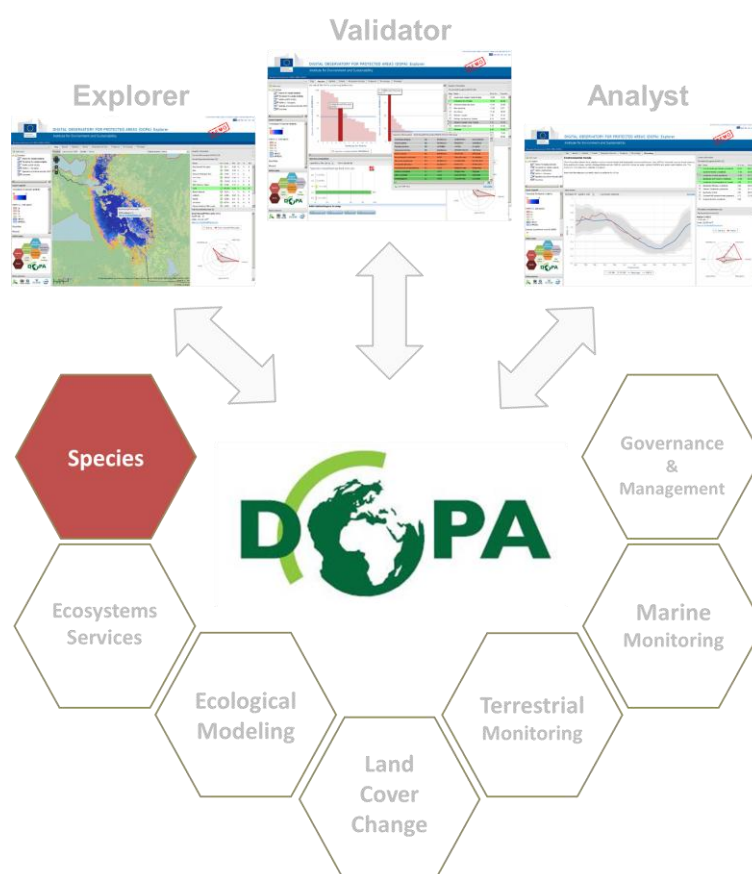


eSpecies

An e-service for mapping species

Andrew Cottam, Michael Schulz, William Temperley, Grégoire Dubois

2013



European Commission
Joint Research Centre
Institute for Environment and Sustainability

Contact information

Grégoire Dubois

Address: Joint Research Centre, Via Enrico Fermi 2749, TP 440, 21027 Ispra (VA), Italy

E-mail: gregoire.dubois@jrc.ec.europa.eu

Tel.: +39 (0)332 786360

Fax: +39 (0)332 789960

<http://dopa.jrc.ec.europa.eu/>

<http://www.jrc.ec.europa.eu/>

This publication is a Technical Report by the Joint Research Centre of the European Commission.

Legal Notice

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Europe Direct is a service to help you find answers to your questions about the European Union
Freephone number (*): 00 800 6 7 8 9 10 11

(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server <http://europa.eu/>.

JRC 87766

EUR 26505 EN

ISBN 978-92-79-35459-5 (pdf)

ISBN 978-92-79-35460-1 (print)

ISSN 1831-9424 (online)

ISSN 1018-5593 (print)

doi: 10.2788/72388

Luxembourg: Publications Office of the European Union, 2013

© European Union, 2013

Reproduction is authorised provided the source is acknowledged.

TABLE OF CONTENTS

Abstract.....	3
Acknowledgements.....	5
Funding	5
1. Introduction	7
2. A Service Oriented Architecture in support to DOPA	9
2.1. eSpecies for indicators on species richness and diversity	10
2.2. Ecosystem services.....	10
2.3. eHabitat for habitat and ecological niche modelling.....	10
2.4. eStation for the monitoring of terrestrial ecosystems	10
2.5. eMarine for the monitoring of marine ecosystems.....	10
2.6. Land cover change and threats to protected areas	10
2.7. Protected areas governance and management effectiveness	11
3. eSpecies	13
3.1. Background	13
3.1.1. Conservation Requirements	13
3.1.2. Audience	14
3.1.3. Sustainability.....	14
3.1.4. Flexibility	14
3.1.5. Performance.....	14
3.2. Analyses	15
3.2.1. Data Sources	15
3.2.2. Products	18
3.3. Information Architecture	26
3.3.1. Design Considerations.....	26
3.3.2. Overall architecture	28
3.3.3. Data Pre-processing	29
3.3.4. Data Management	29
3.3.5. Data Delivery and Web Services (SOA)	31
3.4. Demonstration clients and tools.....	34
3.4.1. Web based tools	34
3.4.2. Integration with Desktop Tools.....	37
3.5. Issues.....	37
3.5.1. Requirements Specification	37
3.5.2. Data Issues	37
3.5.3. Collaborative Development	38

3.6.	Next steps	38
3.6.1.	Improving data quality	38
3.6.2.	Additional Biodiversity Value Analyses	39
3.6.3.	Complementarity Support	39
3.6.4.	Biodiversity Impact Assessment	39
4.	Help and Feedback.....	40
4.1.	DOPA Web site.....	40
4.2.	Engaging with us through the Wiki	40
4.3.	In last resource.....	40
	References	41

Abstract

The Digital Observatory for Protected Areas (DOPA) is conceived as a set of distributed Critical Biodiversity Informatics Infrastructures (databases, web modelling services, broadcasting services, ...) combined with interoperable web services to provide a large variety of end-users including park managers, decision-makers and researchers with means to assess, monitor and possibly forecast the state and pressure of protected areas at local, regional and global scale.

In particular, the DOPA aims to

- provide the best available material (data, indicators, models) agreed on by contributing institutions which can serve for establishing baselines for research and reporting (i.e. Protected Planet Report, National Biodiversity Strategies and Action Plans, ...);
- provide free analytical tools to support the discovery, access, exchange and execution of web services (databases and modelling) designed to generate the best available material but also for research purposes, decision making and capacity building activities for conservation;
- provide an interoperable and, as much as possible, open source framework to allow institutions to get their own means to assess, monitor and forecast the state and pressure of protected areas and help these to further engage with the organizations hosting critical biodiversity informatics infrastructures.

It is the purpose of this document to introduce the readers to eSpecies, the component of the DOPA providing the services focusing on the delivery of species information and products on a 1 km grid at the global scale. In particular, the readers will find here the necessary instruction to access and use our services as well as some information about the possible uses and limitations of the proposed products and services.

Keywords: species, indicators, biodiversity, web services, OGC, Digital Observatory for Protected Areas (DOPA)

Acknowledgements

The following people have also contributed to and assisted in the development of eSpecies by providing invaluable advice: Vineet Katariya (IUCN), Craig Mills (UNEP-WCMC), Tim Robertson (GBIF) and Lars Francke (Cloudera).

Alexandre Tournier, Monica Merlotti, Christian Zanardi and Damiano Binda are the people who have ensure the running and maintenance of the IT infrastructure hosted at EC-JRC that is supporting the DOPA.

Core data sets and web services used in the DOPA have also been made available by the following organizations: UNEP-WCMC, IUCN, GBIF and BirdLife International.

Funding

The DOPA has also been partly supported by the following projects from the European Commission:

The institutional MONDE project of the Land Resources Management Unit, Institute for Environment Sustainability, Joint Research Centre of the European Commission.

BIOPAMA, (<http://www.biopama.org/>) funded by a European Development Fund (DG DEVCO) (2012-2015)

EuroGEOSS (<http://www.eurogeoss.eu/>) funded by the DG RTD of the European Commission (2009-2012)

UncertWEB (<http://www.uncertweb.org/>) funded by the DG INFSO of the European Commission (2010-2013)

PacsBIO, funded by an EC Budget support programme from DG DEVCO to Biodiversity and Protected Areas (2013-2016)

1. Introduction

Protected areas play a key role in conservation programs and in the sustainable use of natural resources. Science-based conservation requires that one has access to a wealth of information on species, ecosystems and threats at the level of the protected area but also at the regional scale to assess priorities, an information that is frequently difficult to access and needs to be regularly verified.

Assessing protected areas for biodiversity conservation at national, regional and international scale, implies that methods and tools are in place to evaluate their physical features such as their proximity to one another, their species assemblages including the frequency and abundance of threatened species, the uniqueness of their ecosystems as well as the threats these areas are exposed to. Typical requirements for such tasks are data on protected areas, information on species distributions and abundance as well as their status on the IUCN Red List of Threatened Species, and information on ecosystems to assess their irreplaceability and monitor changes. By integrating all these data consistently in various indicators, protected areas can not only be evaluated individually but also contrasted against each other for setting conservation priorities. Given the huge amount of information potentially available, information systems need to be developed to ease the processes of collecting, preparing and integrating the data required by the computation of the indicators.

The Digital Observatory for Protected Areas (DOPA) (Dubois et al., 2013) has been developed in collaboration with the UNEP World Conservation Monitoring Centre (UNEP-WCMC), the International Union for Conservation of Nature (IUCN), the Global Biodiversity Information Facility (GBIF), and BirdLife International to support the European Union's efforts *"to substantially strengthen the effectiveness of international governance for biodiversity and ecosystem services (EC/COM/2006/0216 final)"* and more generally for *"strengthening the capacity to mobilize and use biodiversity data, information and forecasts so that they are readily accessible to policymakers, managers, experts and other users"* (UNEP/CBD/COP/10/27).

Conceived around a set of interacting Critical Biodiversity Informatics Infrastructures (databases, web modelling services, broadcasting services, ...) hosted at different institutions, the DOPA is designed to provide to a large variety of end-users, ranging from park managers, funding agencies to researchers, with means to assess, monitor and possibly forecast the state and pressure of protected areas at the local, national and global scales.

In contrast to the previous efforts where most of the data was collected only once and then processed to generate a static set of indicators published on a web site sets (see e.g. Hartley et al., 2007), the DOPA is built around a set of interoperable web services hosted at different institutions. This architecture greatly eases the overall update of the selected data sets and indicators and allows developers to propose an almost infinite number of web based tools for different end-users (Figure 1).

This strategy is encouraged by GEOSS, the Global Earth Observation System of Systems, and its biodiversity component, GEO-BON (Global Earth Observation Biodiversity Observation Network), which have been put in place to better coordinate the efforts to improve and streamline information systems.

Among the main recommendations made by these initiatives, the most commonly encountered when setting up infrastructures involving the exchange, processing and modelling of data are that data should be

- 1) managed as close as possible to its source;
- 2) collected once and documented to allow their use for many purposes;

- 3) easily retrievable and accessible by others;
- 4) interoperable at the syntactic and semantic level to allow their combination for multiple purposes;
- 5) scalable, when applicable, to match other scales;
- 6) shared and, possibly, processed through common, free open-source software tools;
- 7) preserved in persistent repositories and accessible for retrieval by future users.

Similarly, the main functions and models used to compute indicators are likely to evolve as well and their update would require that these can be easily understood and tested thoroughly. This can be best achieved by adopting an interoperable, open source, development framework aiming to put in place a number of independent but interacting components (models, databases, visualization tools ...).

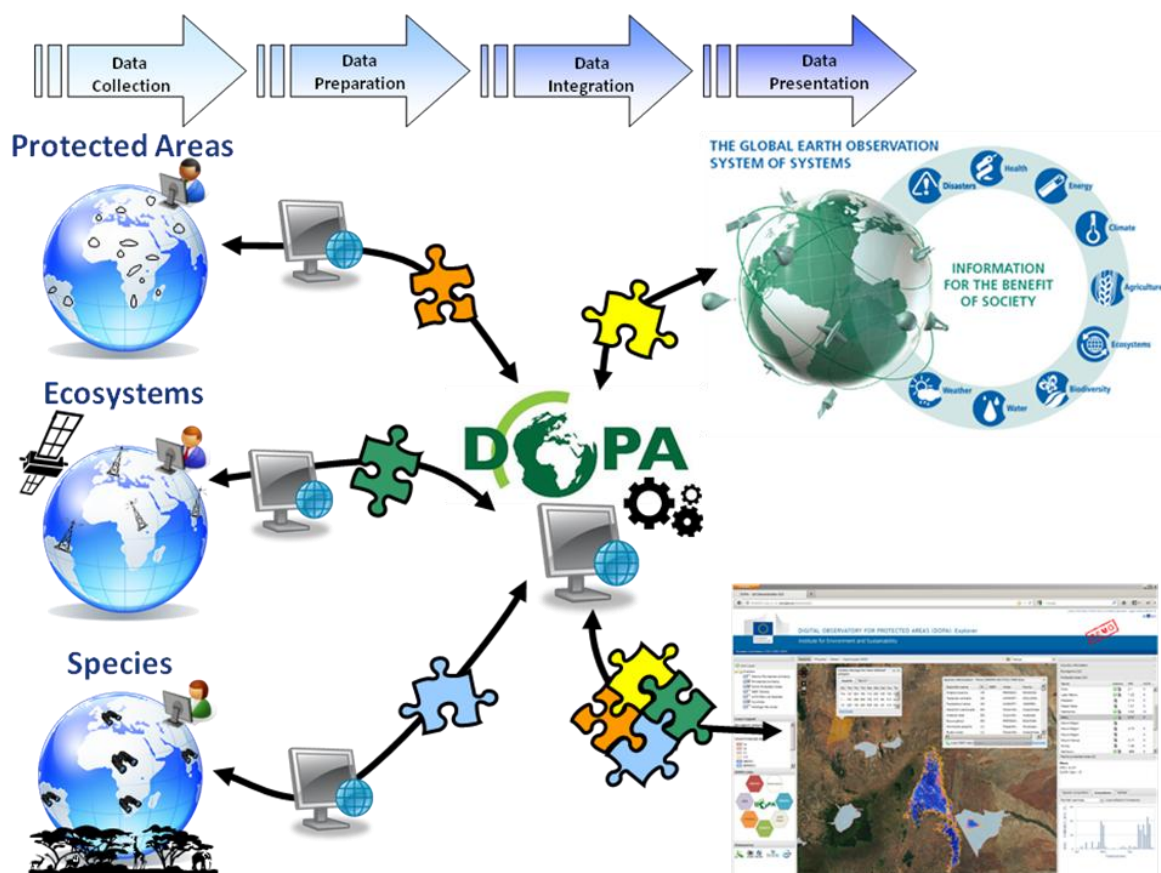


FIGURE 1. FROM GROUND BASED AND REMOTE SENSING OBSERVATIONS TO ENVIRONMENTAL INDICATORS: DATA NEED TO BE COLLECTED, PROCESSED AND PREPARED TO ALLOW THEIR COMBINED USE AND INTEGRATION

2. A Service Oriented Architecture in support to DOPA

The Service Oriented Architecture (SOA) adopted for developing the DOPA is relying on a few institutions that have the mandate to maintain a number of essential databases and services (databases, web modelling services, broadcasting services, etc.). By relying mainly on the following critical infrastructures, the DOPA is stimulating as much as possible a culture of “quality control” for robust science through the whole data process: from the harvesting of the data to their mixing with other sources by different experts when generating new information. The organisation of the DOPA around a web based distributed computing technology should further ease the maintenance and processing of the information. End-users will require only an access to the internet and a web browser to access millions of records, run models and always access the latest information that is available. Similarly, the same infrastructure will allow end-users to contribute with their own information and knowledge to the global information.

For various technical, scientific and even managerial reasons, the development of the DOPA was articulated around 7 core data and model services (Figure 2). The scopes of these services will be summarized in the following sections and advanced users and software developers will find at the following address a more detailed list of the web services underpinning the DOPA: <http://dopa-services.jrc.ec.europa.eu/services/>. This directory of services is continually updated.

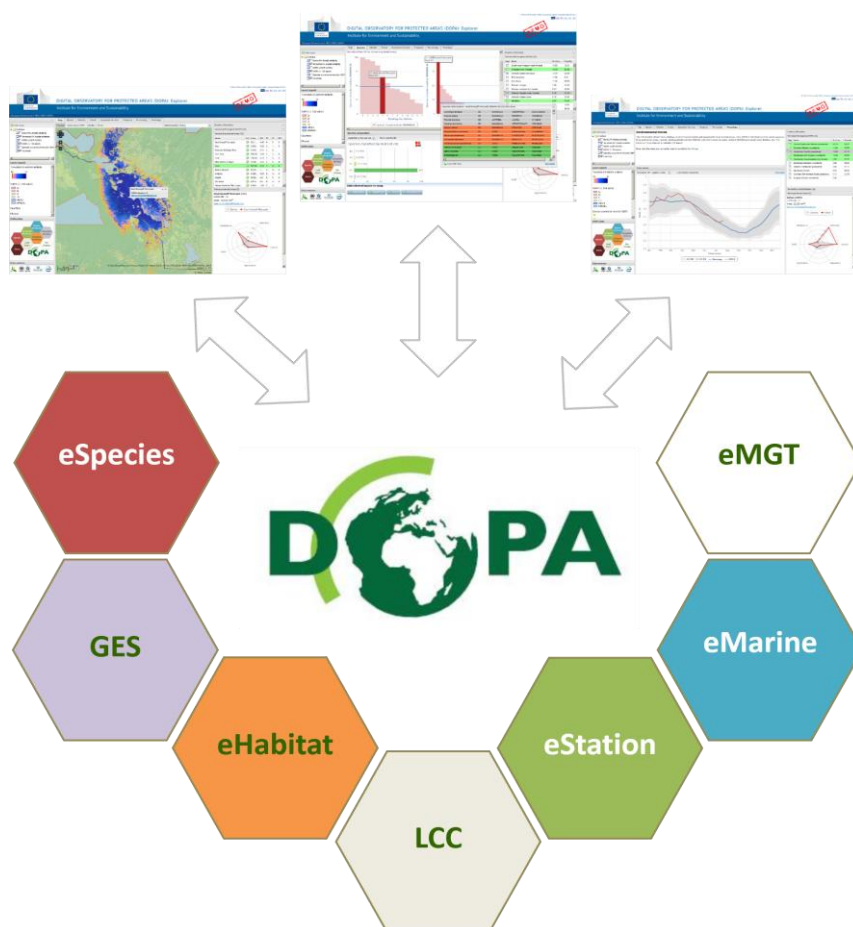


FIGURE 2. ORGANIZATION OF THE CORE DATA AND MODEL SERVICES SUPPORTING THE DOPA

2.1. eSpecies for indicators on species richness and diversity

eSpecies is mainly conceived as a node of the DOPA to process species data, generally hosted by other key institutions such as BirdLife International, GBIF or IUCN, in view to compute a number of indicators on species compositions, species richness and irreplaceability. Because it is the main purpose of this document to provide further details of the services behind eSpecies, we will not discuss it further here.

2.2. Ecosystem services

The mapping of Global Ecosystem Services (**GES**) is an initiative of the JRC to supply maps of Ecosystem Services provided by different types of ecosystems across various spatial scales (Maes et al., 2012). The “benefits that humans derive from ecosystems”, as ecosystem services are defined, support human societies globally through food and water provision, regulation of water flows, use of natural areas for recreation etc. In support to the Ecosystem Service Partnership (<http://www.es-partnership.org/>), this mapping service is focusing in the DOPA on the main ecosystem services provided by protected areas. The Ecosystem Service Mapping service will further provide end-users with quantitative information on an ecosystem service of interest for a specific protected area, its surroundings, or an entire region.

2.3. eHabitat for habitat and ecological niche modelling

eHabitat is conceived as a Web Processing Service (WPS) for computing the likelihood of finding ecosystems with similar properties. A variety of web clients have been developed for different end-users to allow for ecological forecasting in protected areas considering different climate change scenarios, for performing ecological niche modelling or for identifying unique habitats. End-users of the WPS can define the thematic layers for input to the model from various sources, including their own ones. These input layers include data ranging from remote sensing data to socio-economic indicators, thus offering a huge potential for multi-disciplinary modelling (see e.g. Skøien et al., 2012; Dubois et al., 2013)

2.4. eStation for the monitoring of terrestrial ecosystems

The **eStation** is a collecting and processing service designed by JRC to automatically deal with the reception, processing, analysis and dissemination of key environmental parameters derived from remotely sensed data. The measurements are obtained from the SPOT/VGT, SEVIRI/MSG and TERRA-AQUA/MODIS Earth Observation systems. In addition to the web processing service, the eStation offers a number of web clients made available to different end-users for computing ad-hoc thematic products and environmental indicators. Focusing on terrestrial ecosystems, all processing steps of the eStation are easily configurable allowing the user to modify the generated environmental indicators and to implement new ones. The eStation exist as a standalone service and has been distributed in 43 African countries (see Clerici et al., 2013).

2.5. eMarine for the monitoring of marine ecosystems

In essence similar to the eStation, **eMarine** is dealing with earth observations for the marine environment. Monitored physical variables are typically the sea surface temperature and bathymetry. Bio-optical variables used are the coefficients of absorption and particulate backscatter, data on chlorophyll concentration as well as the surface productive layer.

2.6. Land cover change and threats to protected areas

Land cover change (**LCC**) is among the main threats to protected areas and the JRC is working on means to quantify these changes in and around protected areas using web based tools. A web based tool for assessing the impact of protection on land cover in and around protected areas is already in

use in a systematic sampling exercise carried out by BirdLife International and RSPB (Beresford et al. 2013). 100m sample boxes are placed at regular intervals across an Important Bird Area and the 20 km buffer zone surrounding it. In the frame of the DOPA, the tool proposed by Bastin et al. (2012) will be integrated in the DOPA Validator to assess pressures on protected areas.

2.7. Protected areas governance and management effectiveness

In its infancy, this service will focus on management effectiveness (**eMG**T) and protected area governance. It will include means for collecting and analysing information from the field on management and governance (see Hockings 2003) and a service for mapping conservation and research activities in protected areas, ranging from NGOs to governments and universities.

3. eSpecies

This section provides more detailed information about the eSpecies component of the DOPA and describes the conservation requirements, systems being developed, the existing products and services as well as the forthcoming developments.

3.1. Background

The aims of the eSpecies component of the DOPA is to focus on the delivery of species information and products. The history of the development of the eSpecies services began with the publication of the IUCN Red List of Threatened Species and the World Database on Protected Areas (WDPA). These two key strategic conservation datasets were made available to the conservation community but it was very difficult to actually analyse and process them because of the size and complexity of the data. Very simple questions like 'What is where?' that required intersecting these datasets were actually very difficult to answer without significant GIS expertise and expensive computer hardware. Much of the value of these datasets was therefore not being recognised simply because the analyses could not be done and even where the analysis had been attempted it had been done as a one-off.

Through the development of the DOPA, the concept of sustainable biodiversity services at the global scale became part of the project. Once the key partnerships with the data providers had been established the analytical methods could be developed. These are described in the next section.

3.1.1. Conservation Requirements

The conservation community is increasingly producing and using information relating to species and biodiversity within their everyday work in nature conservation. This information comes from a wide variety of sources and is used in many aspects of conservation from local-level conservation management to national level policy support. There are many requirements for information from the different actors involved and although there are some themes in common, no two requirements are ever the same.

The following is a list of the high level conservation information requirements, many of which will be supported within eSpecies (the actual requirements that eSpecies will be delivering against are part of the DOPA Vision Document (Dubois *et al.*, 2013). Examples of some of the eSpecies information products that meet some of these high-level requirements are given in later sections.

Valuation of biodiversity. The assessment and quantification of the biodiversity value for specific geographic regions or places (e.g. protected areas).

Monitoring of biodiversity. The monitoring and surveillance of species and habitats to detect changes to the populations, ranges or statuses.

Systematic conservation planning. The design of protected areas networks and strategies to maximize the protection of biodiversity.

Ecosystem service valuation. Many ecosystem services rely on biodiversity and therefore valuations depend on biodiversity data, e.g. tourism, fisheries management.

Protected Area Management. Local-scale biodiversity data to support the protection and preservation of species and habitats through conservation management.

Environmental Impact Assessment. Multi-scale biodiversity data to support the assessment of the impact on biodiversity through environmental policies or local planning decisions.

Policy Support for the Multinational Environmental Agreements (MEA). Many of the MEAs have statutory reporting obligations that require a range of biodiversity data.

Support for the Convention on Biological Diversity. The main policy instrument for protecting biodiversity requires information mostly at the national level.

The requirements that eSpecies will support will become a part of the wider landscape of key biodiversity services that are being developed and coordinated through the initiatives of the Global Biodiversity Informatics Outlook (Hobern *et al.*, 2013) and the GEO BON initiative¹.

3.1.2. Audience

The main user-community for the eSpecies products and services will be the policy advisors and decision makers within the European Commission, but many of the analyses will be used by conservation organisations and local-level staff. **All of the eSpecies products and services are available in the public domain.**

3.1.3. Sustainability

Another requirement of the eSpecies products, services and tools (and of the DOPA in general) is that they should be produced sustainably. There are a number of aspects to this sustainability in terms of the methodology, the personnel and the technologies.

Firstly, in terms of the methodology the products and services themselves should be easy to update when the constituent data changes. For many of the eSpecies products they are based on datasets that are managed by other organisations who update their data on a regular basis. Therefore, the analyses should be as automated as possible. In addition, some of these datasets are very large in size so any processing should be based on change-only updates if possible.

A second requirement for the sustainability relates to the hand-over of the necessary skills that are required in order to maintain and support the systems in the long term. Many of these sustainability requirements will be described in generic European Commission policy, but the most important aspects for the eSpecies services are that the systems are well documented and use technologies that have widely available skills.

The technological sustainability relates to how the systems have been developed and for the eSpecies this means that there is no reliance on expensive hardware and the systems use open-source tools and technologies wherever possible.

3.1.4. Flexibility

One of the main driving factors behind the development of the eSpecies services has been the ability to deliver products, services and tools that meet a large range of conservation requirements. The conservation community have a diverse set of requirements that change according to the new or emerging threats and these need to be supported with targeted and relevant information. What this means in practice is that instead of designing specific end user tools (for vertical markets) the approach has been to design small interoperable information services that can be assembled in many different ways, a bit like a Lego™ toy where individual components are offering an endless number of combinations. These separate building blocks can then be brought together to create targeted tools and niche products for specific end-users - and these don't have to be developed by JRC but could well be developed by other organisations. The Service Oriented Architecture (SOA) supports this style of systems development.

3.1.5. Performance

Global information on biodiversity is collected at many geographic scales and in many different thematic areas, such as species occurrences, land cover and habitat maps etc. and much of this

¹ <http://www.earthobservations.org/geobon.shtml>

information is managed and delivered locally. The volume, diversity and complexity of the data are all increasing rapidly and conventional information systems and databases are at the limits of their capabilities in retrieving and delivering this information for specific uses.

The eSpecies systems will be based on a new family of technologies that have been developed to cope with this significant increase in data volume. These technologies have been called 'big data' technologies and are characterized by their ability to be able to store and analyse petabytes of information on cheap consumer hardware. The hardware is able to scale to increases in the volume of data. This architecture means that the eSpecies systems will be able to support local and global scale analyses, some in real time or near-real time. The information architecture for the eSpecies systems is described in more detail in later chapters.

3.2. Analyses

The eSpecies analyses have been developed to support a wide range of requirements as outlined in earlier sections. These analyses have been based on datasets provided by partners organisations and JRC have developed value-added products on top of these constituent datasets. If the constituent datasets are updated, then these downstream products and services can be automatically updated and delivered back out to the conservation community. This section describes these value-added analyses and the constituent datasets.

3.2.1. Data Sources

The analyses that are part of the eSpecies products and services are based principally on two main datasets: the World Database of Protected Areas and the Red List of Threatened Species. There are other datasets that are used for some analyses but most are based on either the WDPA or the RL. These are described below.

World Database of Protected Areas (WDPA)

The WDPA² is a knowledge product that is owned by the International Union for the Conservation of Nature (IUCN) and the UNEP-World Conservation Monitoring Centre (WCMC).

The database is the most comprehensive and authoritative source of information on global protected areas and it has been developed over the last 30 years. The database contains spatial information on protected areas throughout the world that have been provided by the national authorities responsible for nature conservation in each country. The spatial information is also supplemented with information on how protected the protected area is (using the IUCN Management Category) and other information such as when it was designated. The database is updated regularly with new information from different countries on a monthly basis and the results are made available through the publically available website called 'Protected Planet' (see Figure 3). JRC have a data sharing agreement with WCMC to be able to use the WDPA for analytical purposes to support the DOPA and to redistribute these analyses back to the community. The WDPA currently contains information on over 200,000 protected areas globally. However, there are still some issues with data quality that can have an impact on derived products (see the data quality section for more information).

² <http://www.wdpa.org/>

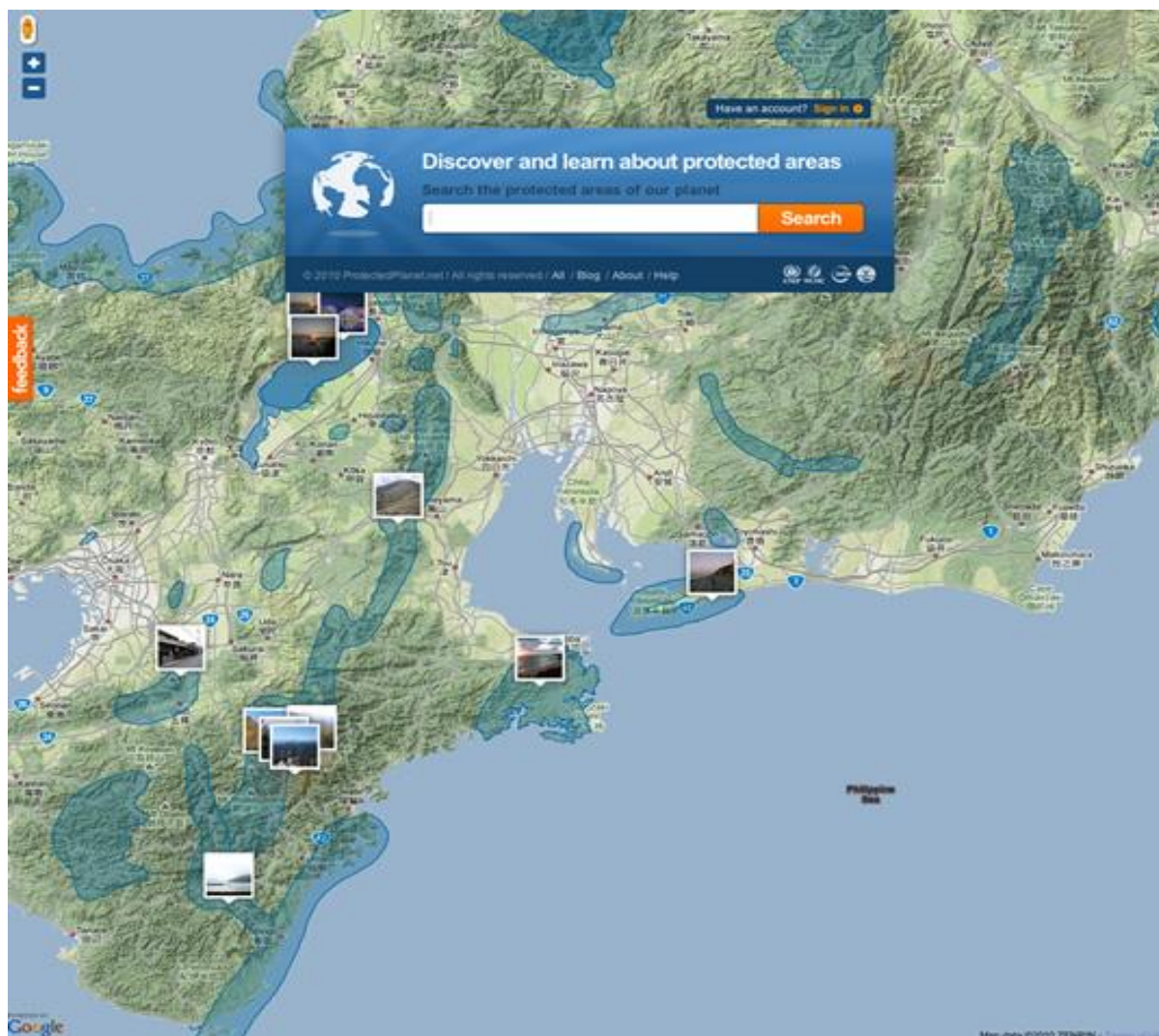


FIGURE 3. SCREEN CAPTURE OF THE UNEP-WCMC PROTECTED PLANET WEBSITE (WWW.PROTECTEDPLANET.NET)

IUCN Red List of Threatened Species

The second major dataset that is used by JRC in their eSpecies analyses is the IUCN Red List of Threatened Species³. This dataset is one of the main knowledge products jointly owned by IUCN, WCMC and BirdLife International and is the most comprehensive and authoritative source on the conservation status of biodiversity.

Currently, the Red List includes assessments for ~50,000 species, with many taxa having been globally assessed, including all mammals, birds, amphibians, freshwater crabs, conifers, and cycads. Distribution range maps for more than 30,000 species are made freely available on the IUCN Red List website (Figure 4), or via partners such as BirdLife International. These range maps represent the Extent of Occurrence of the species and these are synonymous with the species historic range in many cases.

The process of red listing species is carried out by the IUCN Red List Specialist Groups and these groups review specific taxonomic groups every few years.

³ <http://www.iucnredlist.org/>

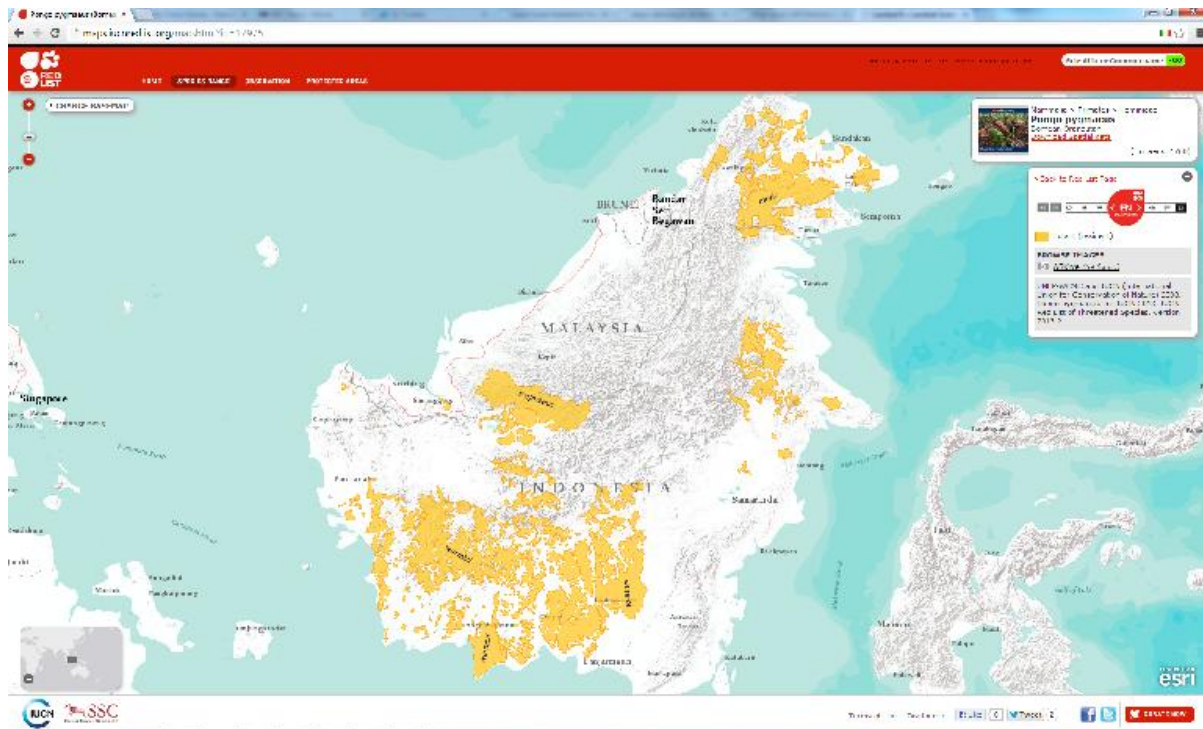


FIGURE 4 SCREEN CAPTURE OF THE IUCN RED LIST OF THREATENED SPECIES WEBSITE. THIS EXAMPLE SHOWS THE RANGE OF THE ORANG UTAN (*PONGO PYGMAEUS*)

The information on each species includes its global conservation status (e.g. critically endangered, endangered, threatened etc.) and information on its particular status at the local level (e.g. whether it is probably present, possibly present etc.). Species are classified by the IUCN Red List into nine groups set through criteria such as rate of decline, population size, area of geographic distribution, and degree of population and distribution fragmentation.

- Extinct (EX) – No known individuals remaining.
- Extinct in the Wild (EW) – Known only to survive in captivity, or as a naturalized population outside its historic range.

Threatened species fall into one of the following three categories

- Critically Endangered (CR) – Extremely high risk of extinction in the wild.
- Endangered (EN) – High risk of extinction in the wild.
- Vulnerable (VU) – High risk of endangerment in the wild.

All other species fall in these last categories

- Near Threatened (NT) – Likely to become endangered in the near future.
- Least Concern (LC) – Lowest risk. Does not qualify for a more at risk category. Widespread and abundant taxa are included in this category.
- Data Deficient (DD) – Not enough data to make an assessment of its risk of extinction.
- Not Evaluated (NE) – Has not yet been evaluated against the criteria.

Species occurrences

In addition to the two principal datasets listed above the eSpecies analyses will use information from the Global Biodiversity Information Facility (GBIF) (see Figure 5). The Global Biodiversity Information Facility (GBIF) Secretariat is facilitating free and open access to species data worldwide via the Internet to underpin sustainable development. GBIF provides currently access to almost 400 million records derived from specimen collections and field observations⁴. Priorities, with an emphasis on promoting participation and working through partners, include mobilizing biodiversity data, developing protocols and standards to ensure scientific integrity and interoperability, building an informatics architecture to allow the interlinking of diverse data types from disparate sources, promoting capacity building and catalysing development of analytical tools for improved decision-making.

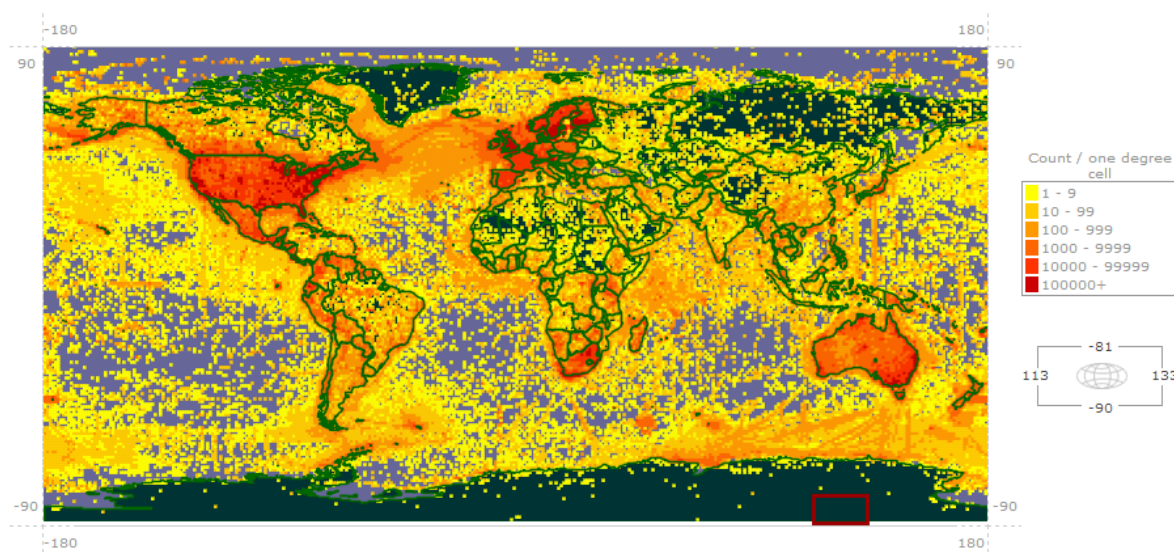


FIGURE 5. THE GLOBAL BIODIVERSITY INFORMATICS FACILITY (GBIF) DATA PORTAL INCLUDES MAPS OF MORE THAN 400,000 SPECIES OCCURRENCE RECORDS. THIS IMAGE SHOWS THE DENSITY OF DATA REGISTERED WITHIN THE GBIF NETWORK INDEX.

Other data

Other reference datasets used by eSpecies include the GAUL Administrative Dataset⁵ and the Worldwide Fund for Nature (WWF) Ecoregion dataset⁶.

3.2.2. Products

There are many definitions for what constitute an information 'product' within the field of biodiversity informatics and the terms can often be misleading. Species range maps, biodiversity indicators, species lists, population trend charts etc. are just some of the many physical outputs that are produced and used within the community for conservation purposes. However, the main emphasis for these products is the fact that they can be used to measure conservation *outcome*, i.e. once a particular conservation goal or target has been set, these products can be used to measure progress against that target. They may be used directly or indirectly - for example, if a conservation goal is to reintroduce a species to a particular protected area then clearly a population trend chart for that species is the ideal product to measure progress against that target. An indirect

⁴ <http://data.gbif.org/welcome.htm>

⁵ <http://www.fao.org/geonetwork/srv/en/metadata.show?currTab=simple&id=12691>

⁶ <http://worldwildlife.org/publications/global-200>

use would be to support overall conservation goals such as the designation of new protected areas based on species distribution information. In this case the species range data is a part of the overall information that is required to designate the new protected area.

The eSpecies products that have been developed within the DOPA framework have focused on the most common requirements - the first of which is the 'what is where?' question. Which species occur at particular locations - whether that is protected areas or countries. The other common question is 'which areas have the highest biodiversity?' These two questions are in the Biodiversity Valuation and Protected Area Management categories given in the section on conservation requirements. Additional products will be developed in the future, but these key analyses have been prioritised. The following sections outline these and other analyses.

The products themselves are all available in the public domain from the DOPA Web Services which are described in the section on Data Delivery.

Biodiversity Valuation

The assessment of biodiversity value is a large subject that encompasses ecology, conservation, ecosystem services, ethics and other factors and there are many different analyses that have been done. The most common measure of biodiversity value is based on the amount of biodiversity (i.e. the total number of species or *species richness*) and also on the constituent value of these species. Many different measures of species richness have been proposed and published that measure species richness within a habitat, between habitats and between regions (known as α , β and γ diversity). Each of them have their relative advantages and disadvantages. For a review of the different methodologies to measure species richness see Brooks *et al.* (2006) and Gotelli *et al.* (2011). Then on top of this simple species richness is the value, or weighting, given to each individual species. This weighting could be based on simple taxonomy, for example large carnivores could get a higher weight because of their position in high trophic levels of the food chain. Other factors could include: the species taxonomic distinctiveness, rarity, range size or commercial value.

There are therefore a high number of potential analyses that could be done simply using the information on the IUCN RL on its own or with other datasets. However, because of the size and complexity of any analyses involving the RL (see the section on data management) many of these analyses had not been attempted - even a simple map of the number of species found at particular locations had not been produced. It was therefore decided to produce two products for biodiversity valuation within eSpecies: a species richness maps at the 1 km resolution for all species and the main taxonomic groups based on the RL; and the species irreplaceability values for protected areas. These analyses involved considerable management and processing challenges that are described in the Data Processing section.

The 1 km Global Species Richness maps

The current (December 2013) 1 km² species richness maps used in eSpecies have been generated for all species and also for the higher level taxonomic groups, e.g. by major taxonomic divisions using the 2011 species range maps from the IUCN Red List. These maps can be used as coarse indicators of overall species richness at the global scale, but these will not be detailed enough for local or regional hotspot analyses. More detailed species monitoring and range data would be need for these types of requirements and this is something that is planned for the future. However, these sorts of outputs are of value to various conservation NGOs who may be taxonomically focused as it highlights those parts of the world that are particularly rich for their species group. Some example of the species richness maps are shown below.

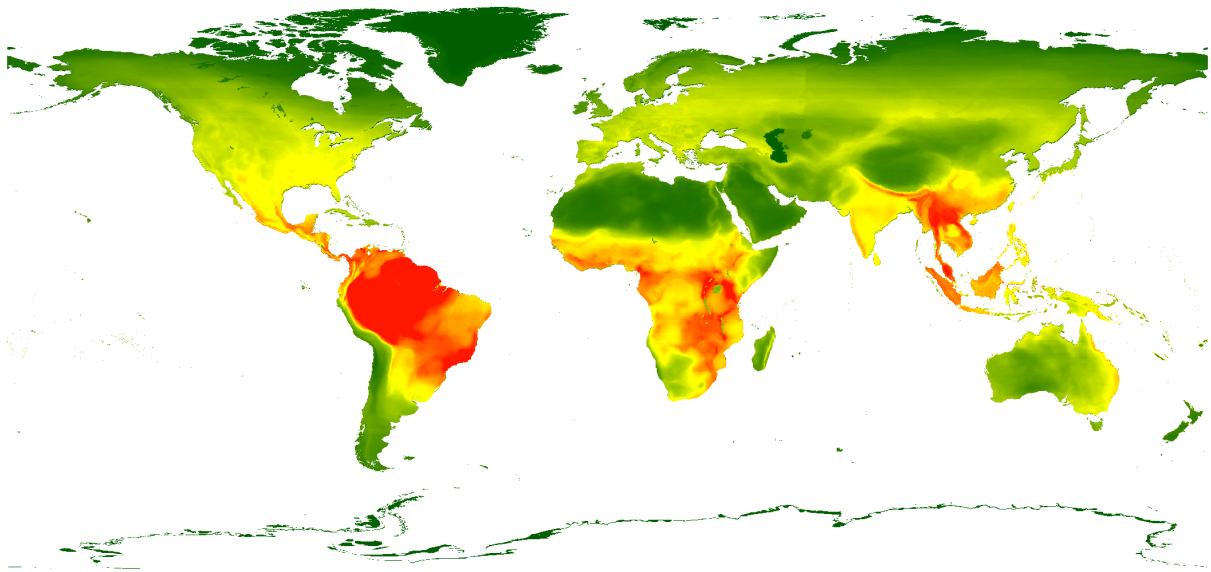


FIGURE 6. EXAMPLE OF A SPECIES RICHNESS MAP FOR ALL SPECIES ON THE IUCN RED LIST OF THREATENED SPECIES. GREEN REPRESENTS LOW SPECIES RICHNESS AND RED HIGH SPECIES RICHNESS.

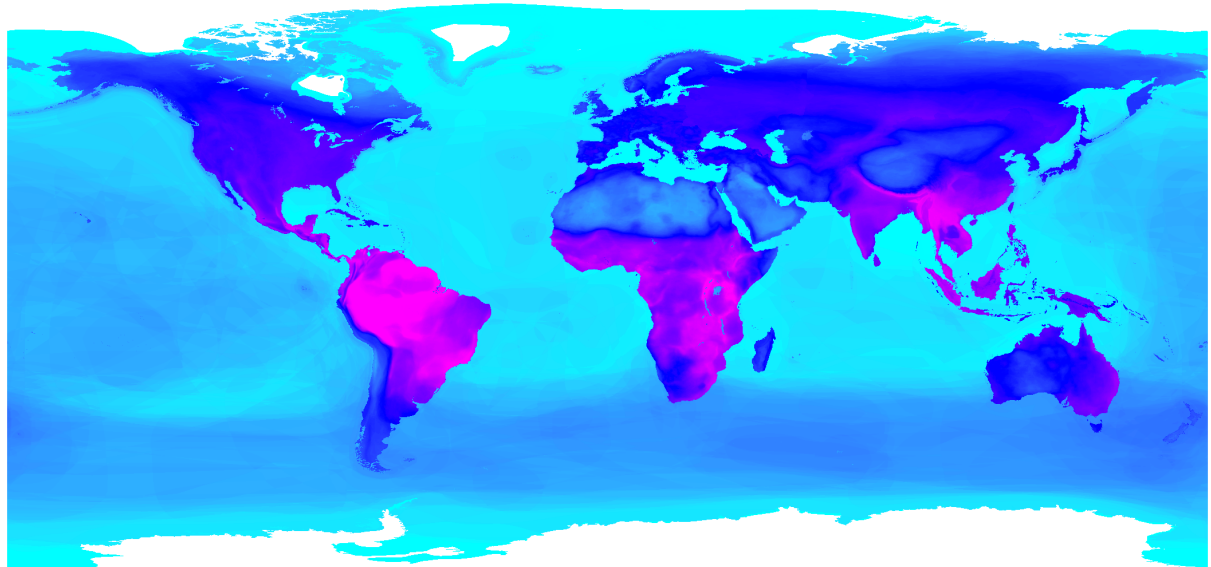


FIGURE 7. EXAMPLE OF A SPECIES RICHNESS MAP FOR BIRD SPECIES BASED ON THE 14,000 BIRD SPECIES ON THE RED LIST. WHITE REPRESENTS LOW SPECIES RICHNESS AND PURPLE HIGH SPECIES RICHNESS.

Species Irreplaceability indicator (SI)

The SI proposed by Hartley *et al.* (2007) is derived here from the 2011 species range maps from the IUCN Red List generated for three taxa (birds, mammals, amphibians) and the WDPA version from May 2013. The species irreplaceability (SI) indicator is calculated for each protected area by counting how many protected areas a species occurs in (n), and adding $1/n$ to the SI index of each of those protected areas. The same procedure was carried out for all species in a given taxon. The higher the value of the SI for a protected area, the higher the number of species found in very few other protected areas and/or the higher the number of endemic species in the PA. In other words, the

higher the SI, the more important is the role of this PA for conserving biodiversity. This approach has the benefit of accounting for the network of protected areas. Any change to the network or the size of the protected areas will impact the SI.

Further normalizing the SI indicators on a scale of 1-100, one can have an idea of the relative conservation value of the protected area for each taxon by means of the radar plot (see Figure 9) or by a bar chart showing the ranking of each indicator of the protected area (see Figure 8). The SI suffers from the limitations indicated in Chapter 2. Species with smaller ranges are more likely to trigger a higher SI and species with large ranges will suffer from the fact that connectivity of protected areas is not taken into account and the critical role of corridors in maintaining viable habitats therefore not considered. There is also a concern that the species maps are sometime not accurate enough to be used in conjunction with small protected areas. Hartley *et al.* (2007) have tried various combinations of species maps and found that the ranking of protected areas based on the SI is robust to changes in the species maps although this observation still needs to be further assessed with a multi-scale analysis of the SIs, from country down to protected area level.

One should note that the SI indicator attributes the same weight to all species independently of their taxon or their threat category on the IUCN Red List of Threatened Species. Because threatened species tend to have smaller distributions, and are therefore found in fewer protected areas, they have a greater effect on the indicator score of the protected area. However, this will still give more emphasize to small endemic species in comparison to larger species which might need to be protected by larger areas and more protected areas, such as rhinoceros and lions.

The species irreplaceability indices have also been calculated for all protected areas lying within specific countries and ecoregions so that comparisons between them at these spatial scales can also be made. These products are shown in the DOPA Explorer (Beta)⁷.

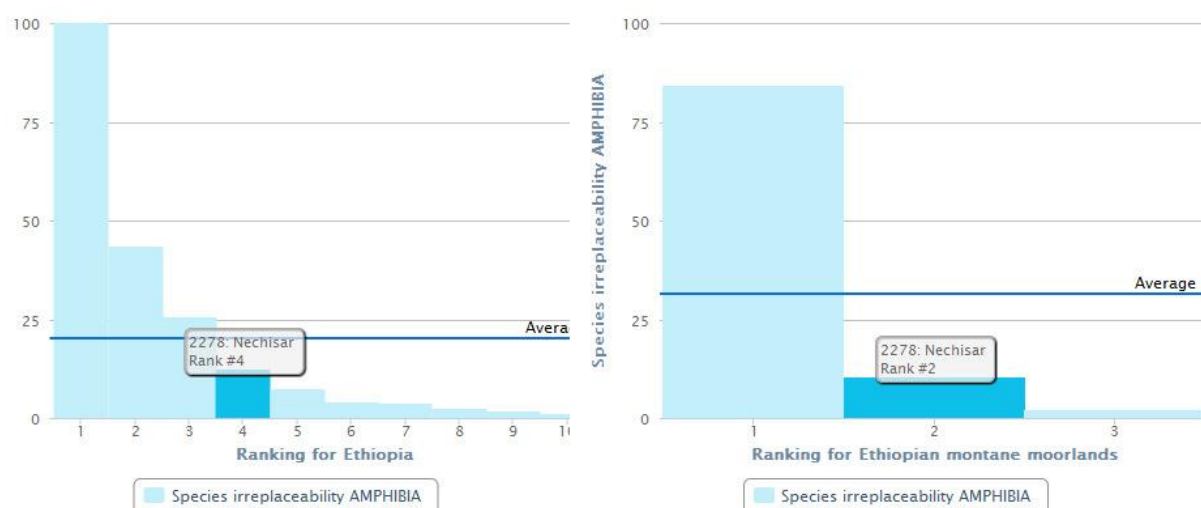


FIGURE 8. BAR CHARTS SHOWING THE RANKING OF PROTECTED AREAS IN ETHIOPIA ACCORDING TO THE VALUES OF THE SPECIES IRREPLACEABILITY (SI) INDICATOR. HERE THE RELATIVE VALUE OF THE SI FOR THE NECHISAR PROTECTED AREA IN ETHIOPIA IS SHOWN AT THE COUNTRY (LEFT) AND ECOREGION (RIGHT) LEVEL.

⁷ <http://dopa.jrc.ec.europa.eu/explorer/>

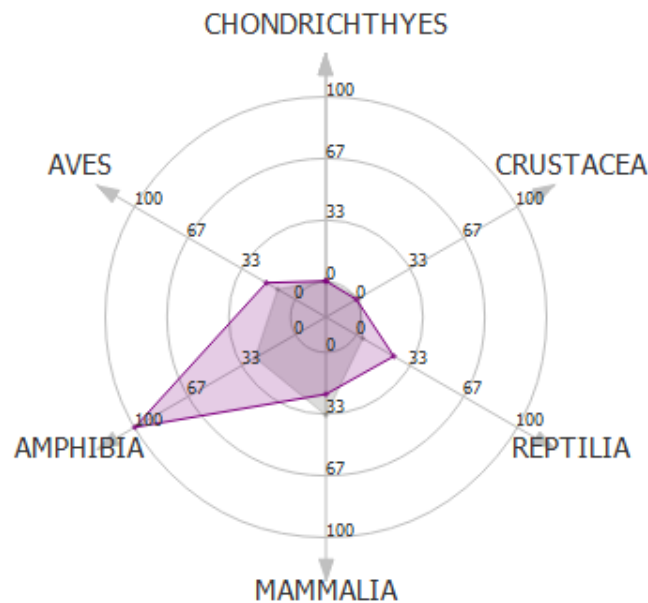


FIGURE 9. RADAR PLOT SHOWING THE SPECIES IRREPLACEABILITY INDICATORS FOR 6 TAXONOMIC GROUPS FOR A PARTICULAR PROTECTED AREA (IN PINK) COMPARED TO A THE NATIONAL AVERAGE OF THE SPECIES IRREPLACEABILITY COMPUTED OVER THE OTHER PROTECTED AREAS. THIS EXAMPLE IS FOR SALONGA NATIONAL PARK IN THE DEMOCRATIC REPUBLIC OF THE CONGO AND SHOWS THAT THIS PARTICULAR PROTECTED AREA IS RICH IN AMPHIBIANS COMPARED TO THE NATIONAL AVERAGE.

Systematic Conservation Planning

Systematic Conservation Planning involves the designation of protected areas based on the principle of maximizing the protection of biodiversity across the whole network of protected areas. This process involves looking at the species that are present within the whole network through *complementarity*, i.e. how the addition of each protected area complements the species that are already protected in other protected areas. It does not work by assessing the biodiversity values of protected areas in isolation.

The eSpecies analysis that contributes to the systematic conservation planning process is described below. In the future new products will be developed that can directly integrate with conservation planning tools, such as the protected area prioritization tool 'Marxan'⁸ (Ball, Possingham and M. Watts, 2009).

Gap Analysis

One of the important products in the complementarity process is the identification of areas where species have no protected - this is also called a gap analysis. A gap analysis has been done on the IUCN RL from Nov 2011 and the protected areas from May 2013 and the results show those areas where species occur with no protection. An additional analysis that is planned is to do a hotspot analysis on these areas to determine the most important areas for unprotected species (and these could be weighted by IUCN status). An example of the gap analysis for the Australasian region is shown below.

⁸ <http://www.uq.edu.au/marxan/>

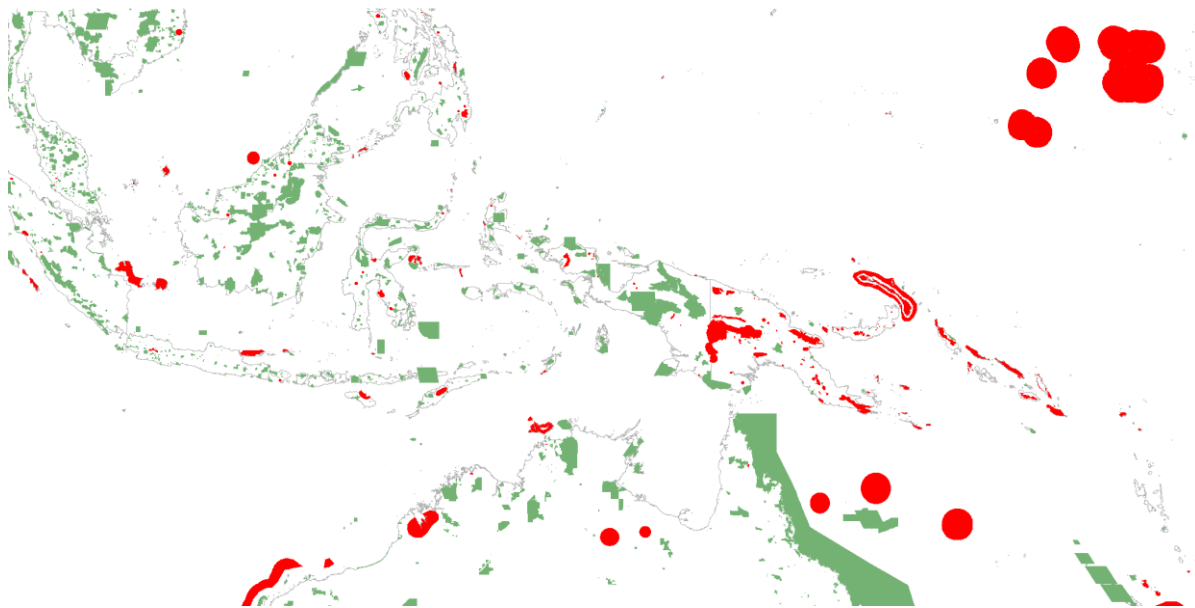


FIGURE 10. GAP ANALYSIS OF SPECIES PROTECTION IN THE AUSTRALASIAN REGION. THE GREEN AREAS ARE PROTECTED AREAS AND THE RED AREAS ARE INDIVIDUAL SPECIES RANGES FOR THOSE SPECIES THAT HAVE NO PROTECTED AREA COVERAGE.

Local-scale products

At the local scale the information requirements for conservation relate to protected area management and environmental impact assessment. Information products on what species occur within protected areas or areas that are threatened with development are key in informing local-scale conservation decisions. Despite the fact that the RL data is in many cases a wide approximation of the actual species range, species lists for protected areas can be useful in determining potential biodiversity and for looking at historic ranges etc. These products can also be useful in local scale planning applications, e.g. for marine and coastal development, in determining if there are globally endangered or threatened species at risk.

The local-scale products would benefit greatly from the improvements in data quality associated with the RL and future planned work includes the DOPA Validator to improve data quality at the local level. Data from GBIF and other biodiversity data clearing houses like Map of Life⁹ will also contribute to better analyses at the local level.

The spatial scale over which the species data is aggregated and summarised is also an important component of the analysis. At this stage, the RL data has been summarised for protected areas and countries but future developments will allow the computation over user-defined geometries. For example, if the geometry of a proposed new road or a pipeline is known it will be possible to get a list of those species whose range intersects with the proposed route.

Species lists for protected areas

The protected area species lists are one of the key products from eSpecies that can be used for a large number of purposes from conservation management to development of tourism. These products can be opened directly in desktop software (for example Microsoft Excel) as the example in Figure 11 shows.

⁹ <http://www.mappinglife.org/>

	A1							
	A	B	C	D	E	F	G	H
	wdpaid	juen_species_id	taxon	common_name	kingdom	phylum	class	order
1	785	17975	Pongo pygmaeus	Bornean Orangutan	Animalia	Chordata	Mammalia	Primates
2	785	44166	Aonyx cinerea	Asian Small-clawed Otter	Animalia	Chordata	Mammalia	Carnivora
3	785	14352	Nasalis larvatus	Proboscis Monkey	Animalia	Chordata	Mammalia	Primates
4	785	9760	Helarctos malayanus	Sun Bear	Animalia	Chordata	Mammalia	Carnivora
5	785	18766	Pteropus vampyrus	Large Flying-fox	Animalia	Chordata	Mammalia	Chiroptera
6	785	12555	Macaca nemestrina	Southern Pig-tailed Macaque	Animalia	Chordata	Mammalia	Primates
7	785	41790	Rusa unicorn	Sambar	Animalia	Chordata	Mammalia	Cetartiodactyla
8	785	22035	Trachypitecus cristatus	Silvery Lutung	Animalia	Chordata	Mammalia	Primates
9	785	41690	Arctictis binturong	Binturong	Animalia	Chordata	Mammalia	Carnivora
10	785	41772	Sus barbatus	Bearded Pig	Animalia	Chordata	Mammalia	Cetartiodactyla
11	785	10551	Hylobates muelleri	Müller's Bornean Gibbon	Animalia	Chordata	Mammalia	Primates
12	785	21488	Tarsius bancanus	Horsfield's Tarsier	Animalia	Chordata	Mammalia	Primates
13	785	12763	Manis javanica	Sunda Pangolin	Animalia	Chordata	Mammalia	Pholidota
14	785	8437	Exilisciurus exilis	Least Pygmy Squirrel	Animalia	Chordata	Mammalia	Rodentia
15	785	19376	Ratufa affinis	Pale Giant Squirrel	Animalia	Chordata	Mammalia	Rodentia
16	785	136603	Neofelis diardi	Sunda Clouded Leopard	Animalia	Chordata	Mammalia	Carnivora
17	785	18128	Presbytis hosei	Hose's Leaf Monkey	Animalia	Chordata	Mammalia	Primates
18	785	41689	Hemigalus derbyanus	Banded Civet	Animalia	Chordata	Mammalia	Carnivora
19	785	10975	Kerivoula intermedia	Small Woolly Bat	Animalia	Chordata	Mammalia	Chiroptera
20	785	12421	Lutra sumatrana	Hairv-nosed Otter	Animalia	Chordata	Mammalia	Carnivora
21								

FIGURE 11. EXAMPLE SPECIES LIST FOR A PROTECTED AREA (KINABALU NATIONAL PARK IN MALAYSIA) ACCESSED DIRECTLY FROM THE DOPA REST SERVICES IN MICROSOFT EXCEL. DESKTOP TOOLS CAN ACCESS THE DOPA SERVICES IN THIS WAY TO MAKE THE PRODUCTS AS EASY-TO-USE AS POSSIBLE.

Species lists for countries

The lists of species by country can be used as one of the inputs into the Convention on Biodiversity (CBD) reporting obligations. The National Biodiversity Strategies and Action Plans (NBSAPs) are one of the main tools at the national level in support to the CBD and identifying the species at risk within each country can be done using these eSpecies products.

Species/Habitat relations

Information products derived from the intersection between the species range data and the Global Land Cover 2000 dataset (and potentially other land cover datasets) can be used to understand species/habitat relationships at range of scales. This information can be used to help with protected area management but it can also be used to help improve the data quality for the RL range data (see Figure 12). See the section on data quality for more information.

Ecosystem Service Valuation

Biodiversity is not only important for the conservation community but also for many other sectors and human kind depends on it for many ecosystem services, including for food, water, fuel and economic revenue. One of the most important potential sources of sustainable revenue for local communities is tourism and the development of tourism for protected areas is a priority in many countries.

Tourism Potential

eSpecies is developing indicators for tourism potential based on the overall attractiveness of the protected areas based on the 'total charisma' of the species that are found within it. This measure of charisma is derived from the total number of photographs of the species found on Flickr - and is therefore a proxy measurement which includes how easy the species is to observe and each picture can be seen as a positive response to an observation. The results of this ranking of species based on charisma are available in the Charismatic Species Gallery (<http://ehabitat-wps.jrc.ec.europa.eu/eSpecies/charisma.html>). A screen shot of such a gallery is shown in Figure 13.

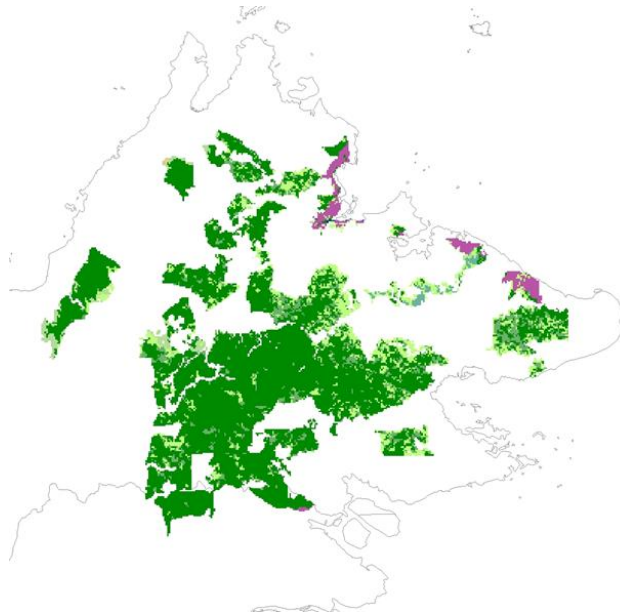


FIGURE 12. SPECIES RANGE DATA FOR ORANG UTAN INTERSECTED WITH THE GLOBAL LAND COVER 2000 DATASET WHICH SHOWS AREAS OF ITS RANGE IN MANGROVE SWAMPS (IN PURPLE) WHICH IT DOES NOT OCCUR IN. MAPS LIKE THIS CAN HELP TO IMPROVE THE IUCN SPECIES RANGE DATA.

Charismatic species gallery x

← → ↻ ehabitat-wps.jrc.ec.europa.eu/eSpecies/charisma.html

Charismatic species gallery

The following species are listed in 'charismatic' order according to the number of images for that species in the [Flickr Image Search API](#) (when searching on "<scientific name>" - only the top rare species. The data for all 30,000 species is available from the DOPA REST Services Directory [here](#), but there is still some work to do on the analysis to control for species range and other factors).

Filter species

Taxonomic group:

Red List Status: ☒ EX ☒ EW ☒ CR ☒ EN ☒ VU ☒ NT ☐ LC ☒ DD











 <p>LION <i>Panthera leo</i> 18574 images</p>	 <p>TIGER <i>Panthera tigris</i> 13733 images</p>	 <p>AFRICAN ELEPHANT <i>Loxodonta africana</i> 8375 images</p>	 <p>LEOPARD <i>Panthera pardus</i> 6737 images</p>	 <p>CHEETAH <i>Acinonyx jubatus</i> 5998 images</p>
 <p>RING-TAILED LEMUR <i>Lemur catta</i> 3666 images</p>	 <p>CHIMPANZEE <i>Pan troglodytes</i> 3159 images</p>	 <p>RED PANDA <i>Ailurus fulgens</i> 3110 images</p>	 <p>HIPPOPOTAMUS <i>Hippopotamus amphibius</i> 2730 images</p>	 <p>WHITE RHINOCEROS <i>Ceratotherium simum</i> 2614 images</p>

FIGURE 13. THE CHARISMATIC SPECIES GALLERY SHOWS THE MOST 'CHARISMATIC' SPECIES ACCORDING TO THE NUMBER OF FLICKR IMAGES FOR THOSE SPECIES AS A PROXY MEASUREMENT.

Other examples

Other examples where species information is helping in the valuation of ecosystem services is in the determination of the extent of mangroves, salt marsh, coral reefs and sea grasses. These extent data can be used to estimate the potential carbon stock.

Research products

In addition to the routine products that are described above which are designed for a wide range for applications, the informatics architecture for the eSpecies is also designed to support the ad-hoc analysis of large datasets (see the information architecture section for more information). These ad-hoc analyses are research-focused and aim to answer a specific question.

One such example was research into the widely held theory of a 'species latitudinal gradient' - that is that you move closer to the equator there are more species. To test this theory the IUCN RL data were rasterized at the 1 km square level and then summarised for each 1 km strip of latitude. This analysis was conducted using a computing cluster based at GBIF and an analysis that would have taken 3-4 months on conventional computers took only 4 minutes. The following chart shows the number of species that occur for each 1 km latitudinal strip - the left edge of the chart is the North Pole. This gradient is shown superimposed over the world in Figure 14)



FIGURE 14. AN EXAMPLE RESEARCH PRODUCT USING INFORMATION FROM THE IUCN RED LIST IS A MAP OF THE OVERALL SPECIES RICHNESS BY LATITUDE. THE MAP SHOWS ON THE RIGHT THE NUMBER OF SPECIES THAT OCCUR WITHIN 1 KM STRIPS WITH INCREASING/DECREASING LATITUDE.

3.3. Information Architecture

The eSpecies and more generally the DOPA products and services aim to support the requirements of a large community of users and this requires a set of robust information systems. Many of these systems are still at an early stage of development, but they will need to be deployed according to a number of design considerations. These are described in the next section.

3.3.1. Design Considerations

Support diverse requirements

The DOPA aims to serve a wide range of users within the conservation community with an equally

wide range of use cases. To support these needs, the overall architecture needs to be able to accommodate a range of analyses and methods. The architecture should make these analyses as easy as possible and should not require the generation of large derived datasets. In addition, requirements will change with time and the architecture will need to be able to cope with new requirements.

Performance and scalability

Many of the datasets that are used in the eSpecies products are global in scale and are derived from highly detailed local-scale data. For example, the RL species range data is based on detailed polygon data from a world coastline dataset. The result of this is a very detailed dataset where species with large ranges, for example many cetacean species, individually represent almost a complete global coastline dataset. This amount of data makes any kind of spatial intersection a very difficult and time-consuming process. In order to meet the performance requirements the architecture must support near real-time querying and analyses and this means sub-second response. For some of the requirements where the spatial dimension is known (for example summarising by country or protected area) these analyses can be prepared in advance and batch processed. However, for the ad-hoc querying, for example returning a list of species for a point, this pre-processing cannot be done easily. Therefore, an architecture that supports rapid real-time querying and analyses is required.

Scalability is also an issue in that it is likely that the volume of data is likely to increase over time and the architecture will need to support this. The current RL includes spatial data on 30,000 species, but it is expected to increase significantly when new taxonomic groups, particularly plants, are assessed.

Support multi-scale analyses

One of the other main requirements is to support spatial analyses at a range of scales. For example, in order for a protected area manager to be able to manage their reserve they may need to know: what species occur within their protected area (local scale analysis); what is the global context for those species (i.e. are some species edge-of-range within the reserve - a global analysis); how connected are these species to other protected areas (regional analysis).

Support cross-dataset analyses

The architecture must also support the ability to conduct cross-sectorial analyses. The most important of these cross-sector analyses will involve datasets that examine threats to biodiversity and protected areas, such as climate change and land cover change. Within the DOPA there are a number of other information services that relate to fire, land-cover, climate etc. which can all be used in any threat assessments at a range of scales.

Support dynamic data

The DOPA is integrating data from a number of different providers including the UNEP-WCMC World Database of Protected Areas (WDPA) and the IUCN Red List. Each of these dataset has its own update mechanism and frequency and the architecture will need to accommodate changes in these source datasets into the derived datasets.

In addition, it is envisaged that there will also be a mechanism allow registered users to update data directly within the eSpecies Services to improve its quality directly. How this information is then validated and passed back to the original data providers also needs to be taken into account.

Use open-source approach

Open data refers to the idea that certain data should be freely available for use and re-use. The European Commission's work in the area of open data is focussing on generating value through re-use of a specific type of data – public sector information, sometimes also referred to as government data. That is all the information that public bodies produce, collect or pay for. This is very much the

case for much of the existing biodiversity information and we adopt the same approach in our technical developments, adopting open source technology. All of the tools and services will be published and made available on an appropriate source-code repository for download and use.

Support to other projects

The open access to our services also encourages their use to support other projects, in particular where the tools or services may need to be extended or customised for local needs. The services of the DOPA are, among others, currently used to support the regional information systems deployed for the BIOPAMA¹⁰ (Biodiversity and Protected Areas Management in the African, Caribbean and Pacific regions) project. While many of the local organisations may already have existing IT systems, connecting these services require the tools and capacities to customize these to be as much accessible as possible.

Service-Oriented Architecture (SOA)

The overall architecture must support a SOA so that the information services can be loosely coupled and combined easily with few dependencies or external requirements. This makes data integration much easier and provides developers (from European Commission and in other institutions and organisations) with the ability to quickly develop solutions that integrate data from the eSpecies services.

Incorporate community-driven data standards

In the same way that the SOA architecture will enable the technical integration of information services, the community-driven data standards will provide the integration of *content*. What this actually means in practice is that the conservation community have a set of agreed data dictionaries (or standards) that are used to describe entities like species or geographies. These data standards are promoted by organisations like GBIF and the Taxonomic Database Working Group and need to be included in the overall architecture.

Support for dataset versioning and metadata

The eSpecies services will need to provide access to the species range data and also to a set of other derived datasets. These will need to be properly documented and version-stamped so that any users of the data can have a full understanding of the sources, process and caveats on the use of the data. They will then be able to establish whether the information is fit-for-purpose. This is particularly important where the data is likely to change regularly (as is the case for the WDPAs). Generic metadata standards will be used wherever possible.

3.3.2. Overall architecture

In order to meet the design considerations set out above, the following architectural components were developed for the eSpecies products and services:

- A pre-processing framework to rasterise, encode and serialise the species range data
- A generic data model based on a simple quad key spatial data structure
- A scalable database technology
- An eSpecies API delivering OGC and REST based services
- A small number of client demonstration applications

¹⁰ <http://www.biopama.org/>

3.3.3. Data Pre-processing

The pre-processing that is done in the eSpecies Services converts the feature data from the RL to simple rows in a 'non-spatial' database. The use of vector data for the species range maps is fast and efficient, particularly for mapping single species, but in order to be able to analyse the data across different species some kind of rasterisation/gridding must be done. This process must also be able to support the analyses from the global to the local scale (in the first instance the local scale needs to be down to 1 km).

The traditional approach would be to rasterise each species range map as a separate discrete dataset and then to use cell-based calculations to derive new products. However, there are three problems with this approach that makes it a poor solution.

Firstly, the available memory. The size of a single species raster would be very large and a huge amount of RAM would be required for the processing (i.e. a global raster at 1 km resolution, e.g. for a cetacean, needs more than 1 Gb of RAM). Combining multiple datasets at the same time to create new outputs would require significant hardware investment. This process could be done using a moving window type approach or incrementally, but this architecture would not support ad hoc spatial querying for a location (since all of the species rasters would need to be loaded in memory at the same time as a stack).

Secondly, the reporting efficiency. Raster datasets can hold attribute data that can be mapped but it is more difficult for them to hold or relate to additional information that may need to be mapped or filtered. It is not possible to join attributes from raster datasets to other non-spatial data and filter the raster accordingly. This reduces the flexibility for reporting and means that new analyses would have to be driven by geoprocessing chains rather than by simple querying.

Finally, storing species range data in a raster format means that all cell values need to have a value even if that value is 'does not occur' in order for the raster dataset to be continuous. This will significantly increase the storage required and reduce the access efficiency, particularly for species with large ranges. What is needed is a mechanism for capturing and managing just the information where the species is present.

The rasterisation process converts the polygon data to raster using the Web Mercator projection (EPSG:900913). Data are then processed using Python to encode the data with unique quad-keys for their 1 km square locations and these 'non-spatial' records are saved to disk and ready to be loaded into a database. This rasterisation/serialisation is done in a multiprocessing environment to make the pre-processing as efficient and as fast as possible (in just over 5 hours with 16 processors for 30 000 species range maps).

3.3.4. Data Management

The data that is produced at the end of the pre-processing step is a simple set of files that refer to individual species ranges which are encoded with quad keys and other attributes. These are described below.

Quad tree addressing

In order to be able to support high performance multi-scale analyses and the ability to query the gridded data by location, some form of spatial hashing within the data is needed. One such model that meets all of the requirements is to use quad tree addressing for the grid cells. This approach is a well-established mechanism for uniquely identifying areas on the earth using a key that is also hierarchical. The inherent hierarchy in quad keys also means that they can be used to dynamically summarise data at a higher spatial scale (Figure 15).

One of the most useful aspects of quad tree addressing is that it provides a spatial framework that is very lightweight - information about an area is a simple key. This key can form the basis of data

exchange, APIs and real-time processing. For example, if UNEP-WCMC could deliver their protected areas information based on the quad key then you could rapidly integrate data on protected areas, species, land cover data etc. into simple web clients. This integration would require very little technological infrastructure or specialist GIS software.

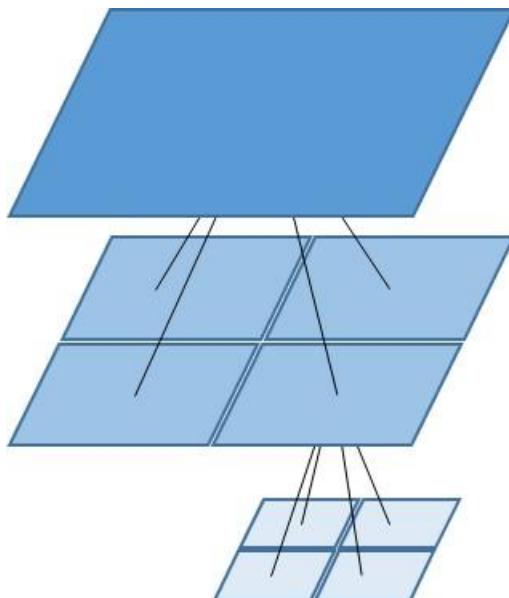


FIGURE 15. AN EXAMPLE OF A QUADTREE STRUCTURE

Database schema

The database schema has been designed to support a wide range of query requirements and to enable cross-sector analyses. The schema is effectively a single large table that holds quad key data corresponding to a particular entity (in most cases a species, but it could be a protected area, country or some other entity). This schema represents a generic data bucket that can be used to manage information on many different types of data, not just species, using the quadkey as the unique spatial location. Data can then be cross-related and analysed using this quadkey.

Database technology

The IUCN RL species range dataset currently contains spatial information on over 30,000 species. Given that the requirements for the analysis are to be able to summarise data at the 1 km square level, then the potential total volume of data is predicted to be of around 80 billion records. The major issue with the database is therefore the performance. Managing such a large number of records in a system that needs to support on-the-fly querying, data update and re-indexing is likely to be a significant challenge. A number of technologies were evaluated for managing these large datasets and the Apache Hadoop system¹¹ was selected as it is a widely used open-source big data system that offers many advantages including scalability. The only disadvantage is that it is reasonably new technology and therefore requires more time in deployment and development.

Data Processing

The generation of products for eSpecies can be done in batch for those datasets, like the RL, that do not change very often. These products can all be processed directly from the Hadoop computing cluster. The cluster also provides excellent processing capabilities for ad-hoc and research based

¹¹ <http://hadoop.apache.org/>

analyses in near-real time (as demonstrated by the species latitudinal gradient analysis). Although these products may be processed and generated on the Hadoop cluster, they may be managed elsewhere for delivery. For example, once the processing of global species richness maps has been completed, the resulting outputs are then published as images on the Web Map Server. The results of non-spatial analyses are stored in Postgresql ready to be published as web services.

3.3.5. Data Delivery and Web Services (SOA)

During the design of the information systems for the DOPA, one of the key considerations was for the analyses to be published as soon as they have been created with very little effort on the part of the data analyst.

Spatial Web Services

All spatial products are available as a set of OGC (Open Geospatial Consortium) standard web services published through an industry standard system called 'Geoserver'. These services are all available from the Geoserver directory here:

<http://Irm-maps.jrc.ec.europa.eu/geoserver/web/>

This directory lists the species richness datasets that have been derived from the RL (Figure 16). All of these species richness datasets have been created at a 1 km spatial scale and are the result of rasterizing all of the species on the RLTS and overlaying all species on top of each other to produce a richness surface. Higher level richness maps are available for birds, amphibians, mammals and reptiles and in addition, the lower taxonomic levels also have richness maps. These services can be consumed in Desktop GIS tools and also incorporated into web sites. For more information on using OGC Web Services see the specific documentation that comes with the specific software.

The screenshot shows the GeoServer Layer Preview web interface. The browser address bar displays the URL: <http://Irm-maps.jrc.ec.europa.eu/geoserver/web/?wicket:bookmarkablePage=:org.geoserver.web.demo.MapPreviewPage>. The page features a sidebar with navigation links: 'About & Status', 'Data', and 'Demos'. The main content area is titled 'Layer Preview' and contains a table listing 153 items (results 1 to 25 shown). The table has columns for 'Type', 'Name', 'Title', 'Common Formats', and 'All Formats'. Each row represents a species richness dataset, with a checkbox in the 'Type' column and a dropdown menu in the 'All Formats' column.

Type	Name	Title	Common Formats	All Formats
<input checked="" type="checkbox"/>	irm-species:ACTINOPTERYGII	ACTINOPTERYGII	OpenLayers KML	Select one
<input checked="" type="checkbox"/>	irm-species:ANIMALIA	ANIMALIA	OpenLayers KML	Select one
<input checked="" type="checkbox"/>	irm-species:LYCOPODIOPHYTA	LYCOPODIOPHYTA	OpenLayers KML	Select one
<input checked="" type="checkbox"/>	irm-species:CRUSTACEA	CRUSTACEA	OpenLayers KML	Select one
<input checked="" type="checkbox"/>	irm-species:REPTILIA	REPTILIA	OpenLayers KML	Select one
<input checked="" type="checkbox"/>	irm-species:RHODOPHYTA	RHODOPHYTA	OpenLayers KML	Select one
<input type="checkbox"/>	irm-species:amazon	amazon	OpenLayers KML GML	Select one
<input checked="" type="checkbox"/>	irm-species:INSECTA	INSECTA	OpenLayers KML	Select one
<input checked="" type="checkbox"/>	irm-species:AMPHIBIA	AMPHIBIA	OpenLayers KML	Select one
<input checked="" type="checkbox"/>	irm-species:GASTROPODA	GASTROPODA	OpenLayers KML	Select one
<input checked="" type="checkbox"/>	irm-species:AVES	AVES	OpenLayers KML	Select one
<input checked="" type="checkbox"/>	irm-species:SELLAGINELLOPSIDA	SELLAGINELLOPSIDA	OpenLayers KML	Select one

FIGURE 16. SCREEN CAPTURE OF THE ONLINE DIRECTORY OF THE SPECIES RICHNESS MAPS AVAILABLE FROM THE GEOSERVER.

Non-spatial Web Services

For non-spatial datasets, a custom web server was developed that allows products that have been created to be published simply by creating a query in the database. These resources are then available from an automatically generated set of web pages that publish the products as 'REST Services'.

REST Services are simply information products that are available to users as a URL - information on what data to retrieve and in which format is all part of the URL. This makes REST Services very easy to consume in many different types of application and tools - enabling the services to get maximum penetration into the conservation community. The services can be used in Desktop tools, e.g. Microsoft Excel, custom web sites and also support delivery of information as SMS, PDF and email. However, in many cases the actual end points to the REST services are intended to be used for computer-to-computer exchange of data, for example in the creation of a web site, rather than for direct reading by a user.

The DOPA REST Services are available from the DOPA REST Services Directory (<http://dopa-services.jrc.ec.europa.eu/services/>) - see Figure 17. This is the main gateway to all the non-spatial web services for the DOPA and new analyses will appear here as and when they are developed and published. Under each main schema, e.g. eSpecies, eHabitat etc. are listed all of the individual Web Services that provide access to a specific piece of information. To access the eSpecies services click on the eSpecies link. It should be emphasised that the services that are published through the directory are intended for computer-to-computer communication and data exchange. Although the service results can be viewed in a browser they are targeted at other *clients* (for example web sites of desktop tools) which can render the results in a better format for viewing and interacting.

Schemas		
Schema	Description	More information
eHabitat	Contains information services relating to habitats. Esp. the habitat replaceability index (HRI) and ecoregion related information.	http://ehabitat.jrc.ec.europa.eu/
eReference	Contains reference services, such a generic administrative unit datasets	http://ereference.jrc.ec.europa.eu/
eSpecies	Contains information services relating to species using information from the IUCN Red List of Threatened Species and other datasets.	http://especies.jrc.ec.europa.eu/
eStation	Contains information services relating to estation. Currently these are limited to fire occurrences.	http://estation.jrc.ec.europa.eu/

FIGURE 17. THE DOPA REST SERVICES DIRECTORY CONTAINS LINKS TO ALL OF THE DOPA NON-SPATIAL WEB SERVICES ([HTTP://DOPA-SERVICES.JRC.EC.EUROPA.EU/SERVICES/](http://dopa-services.jrc.ec.europa.eu/services/))

For example, to get a list of species that occur for a particular protected area, click on the eSpecies link and then on the `get_pa_species_list` link. This brings you to a page which describes the Web Service, its parameters and the information that is returned. Sample links are given at the bottom of the Web Service page.

An example of the JSON format which is returned in response to a request is shown in Figure 18.

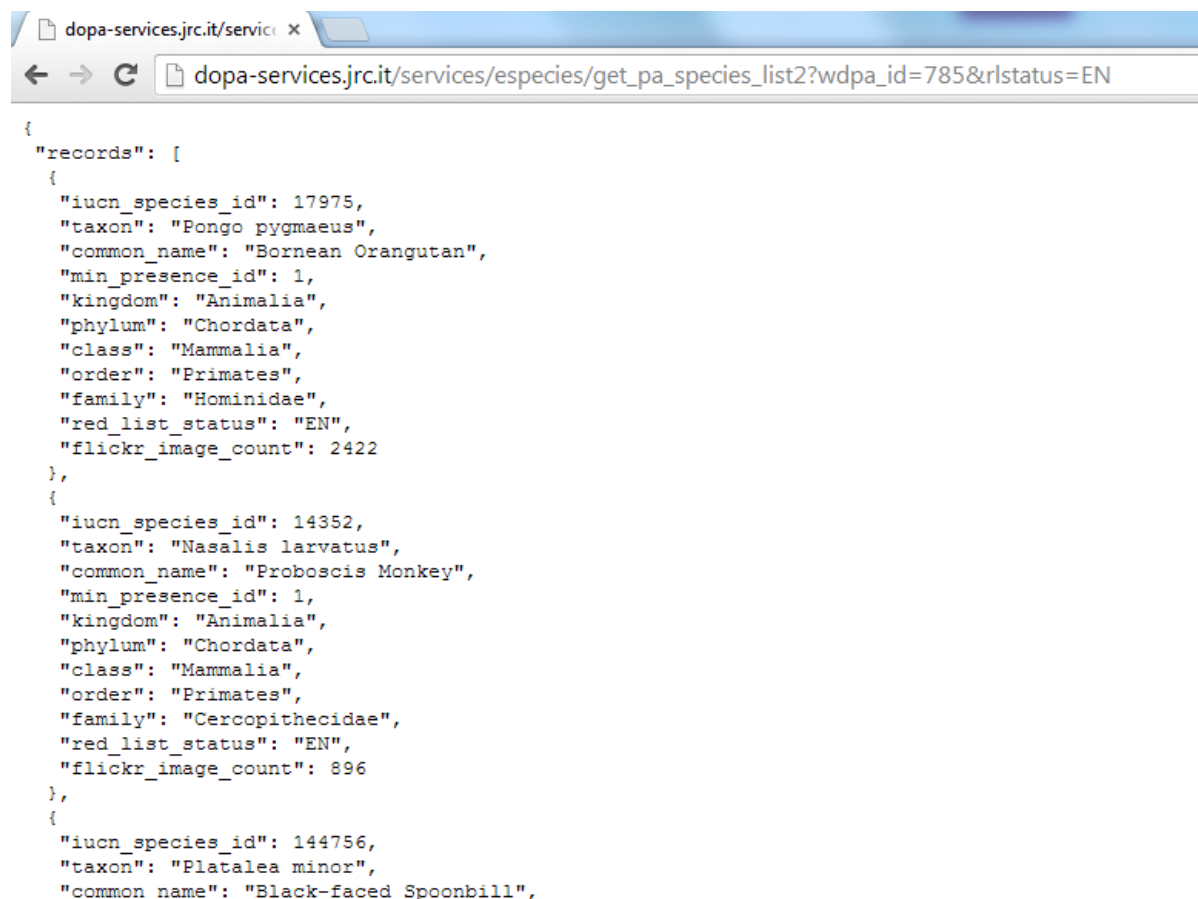


FIGURE 18. EXAMPLE RESPONSE FROM A TYPICAL REST SERVICE. THIS EXAMPLE SHOWS A LIST OF SPECIES FOR A PROTECTED AREA AND THIS INFORMATION CAN BE USED BY A WIDE RANGE OF COMPUTING CLIENTS (E.G. WEB SITES, DESKTOP TOOLS, GIS TOOLS, MOBILE PHONES ETC.)

For full information on how to use the DOPA REST Services Directory see the Help documentation (<http://dopa-services.jrc.ec.europa.eu/services/help>). This includes details of how to use the Web Services in common Desktop tools, like Microsoft Excel.

The eSpecies Web Services are broadly divided into a number of thematic areas according to the following sections. The actual list of services that may be seen in the Services Directory may be different to those that are described below and this is because the services that are made publically available will change with time. It is anticipated that more eSpecies services will be made available as the DOPA develops. The following sections provide links to some of the most important REST services for eSpecies - the description of the products is given in an earlier section.

Species lists

The following services return information about species that occur in geographic areas:

http://dopa-services.jrc.ec.europa.eu/services/especies/get_country_species_list

http://dopa-services.jrc.ec.europa.eu/services/especies/get_pa_species_list

There are also simple services that retrieve the count of species within particular countries or protected areas.

Species Irreplaceability

The following services return information on SI:

http://dopa-services.jrc.ec.europa.eu/services/especies/get_species_irreplacibility_country

http://dopa-services.jrc.ec.europa.eu/services/especies/get_species_irreplacibility_ecoregion

Species charisma

The following services return information about species charisma:

http://dopa-services.jrc.ec.europa.eu/services/especies/get_species_charisma

http://dopa-services.jrc.ec.europa.eu/services/especies/get_species_charisma_all

http://dopa-services.jrc.ec.europa.eu/services/especies/get_tourism_ranks_global

Unprotected species

The following services returns information about the unprotected species:

http://dopa-services.jrc.ec.europa.eu/services/especies/get_unprotected_species

3.4. Demonstration clients and tools

Much of the development work within eSpecies has been on the development of the back-end analytical systems and delivery mechanisms to provide access to the eSpecies products as Web Services. Although these services can be used directly, it is only when the separate information products are combined into useful and targeted tools that they become much more useful. A small number of demonstration clients and tools have been developed for eSpecies and these are described below.

3.4.1. Web based tools

There are a small number of web-based tools that demonstrate the use of eSpecies (and other DOPA) services for specific needs.

DOPA Explorer

The DOPA Explorer tool is a 'shop window' onto the available information products and services that are available for protected areas at the global scale. The current version (Explorer Beta) shows these products in a set of thematic tabs for all protected areas in the world ≥ 150 km² in size (Figure 19). Available in English, French, Spanish and Portuguese, DOPA Explorer (Beta) is a first web based assessment tool where information on 9 000 protected areas covering almost 90% of the global protected surface has been processed automatically to generate a set of indicators on ecosystems, climate, phenology, species, ecosystem services and pressures. DOPA Explorer can so help identify the protected areas with most unique ecosystems and species and assess the pressures they are exposed to because of human development. Ecological data derived from and near real-time earth observations are also made available for the African continent. Inversely, DOPA Explorer indirectly highlights the protected areas for which the information is incomplete.

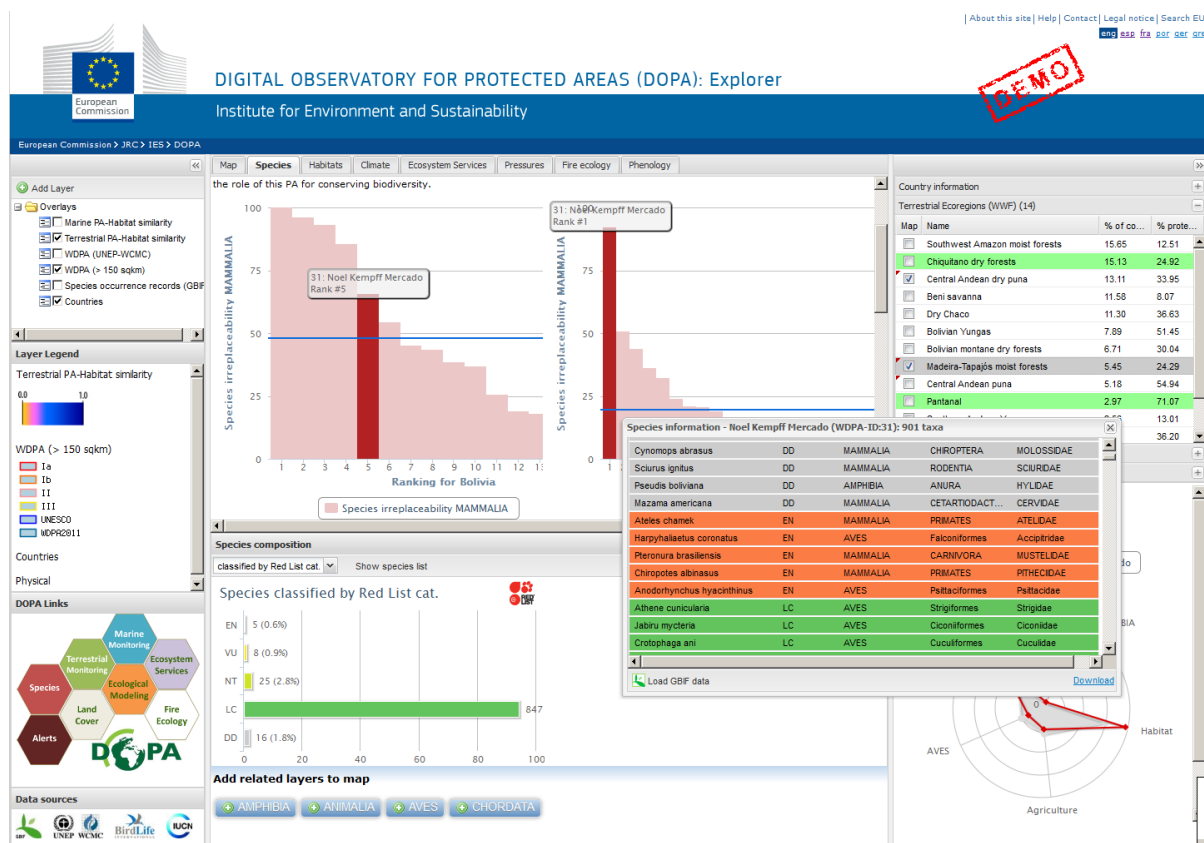


FIGURE 19. THE DOPA EXPLORER TOOL WHICH IS INTENDED AS A DEMONSTRATION CLIENT WHICH INTEGRATES MANY OF THE SPATIAL AND NON-SPATIAL DOPA WEB SERVICES INTO A PROTECTED AREA-FOCUSED WEBSITE.

Protected Areas Species Checker

The protected areas species checker is another web tool that shows information that is being provided by the eSpecies services - this time in a very different format. Here, the species that occur within a protected area are shown together with Flickr images. This particular web tool is designed to provide a feedback mechanism to be able to flag those species that do not occur at a particular protected areas (despite the protected area/species intersection analysis predicting them to occur). The tool will be used by protected area managers and other local actors to help improve the data quality for the RL TS and is an early version of the species component of the DOPA Validator (in prep). An illustration of the current version of the species checker is shown in Figure 20.

The protected area species checker is available at the following address:

<http://ehabitat-wps.jrc.ec.europa.eu/eSpecies/speciesValidator.html?wdpaid=116275>

Global Tourism Potential

Still in development, the global tourism potential tool is used to show those protected areas that have the greatest tourism potential globally based on the species that occur there and on the species charisma. An example showing the protected areas in Indonesia with greatest tourism potential is shown in Figure 21.

The global tourism potential tool is available at the following address:

<http://ehabitat-wps.jrc.ec.europa.eu/eSpecies/tourismGlobal.html>

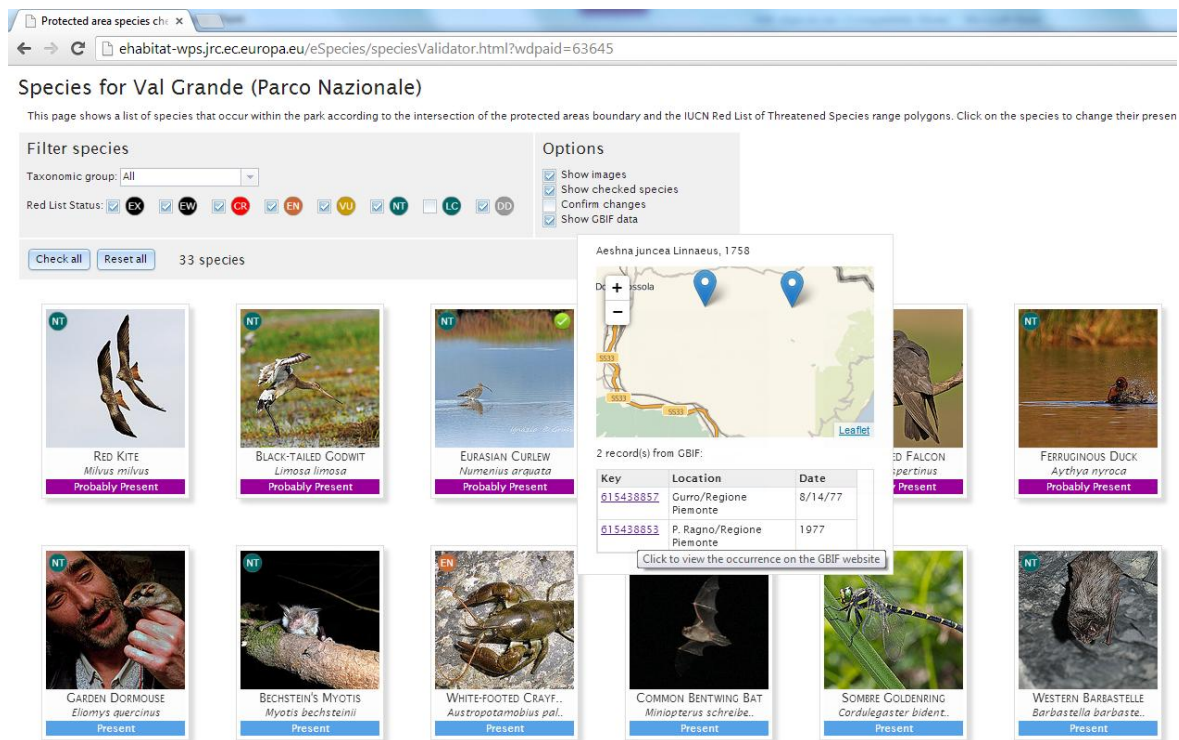


FIGURE 20. THE SPECIES VALIDATOR TOOL IS ANOTHER CLIENT FOR THE DOPA REST SERVICES THAT WILL PROVIDE A MECHANISM FOR PROTECTED AREA MANAGERS AND OTHER LOCAL STAFF TO VALIDATE SPECIES LISTS FOR THE PROTECTED AREAS.

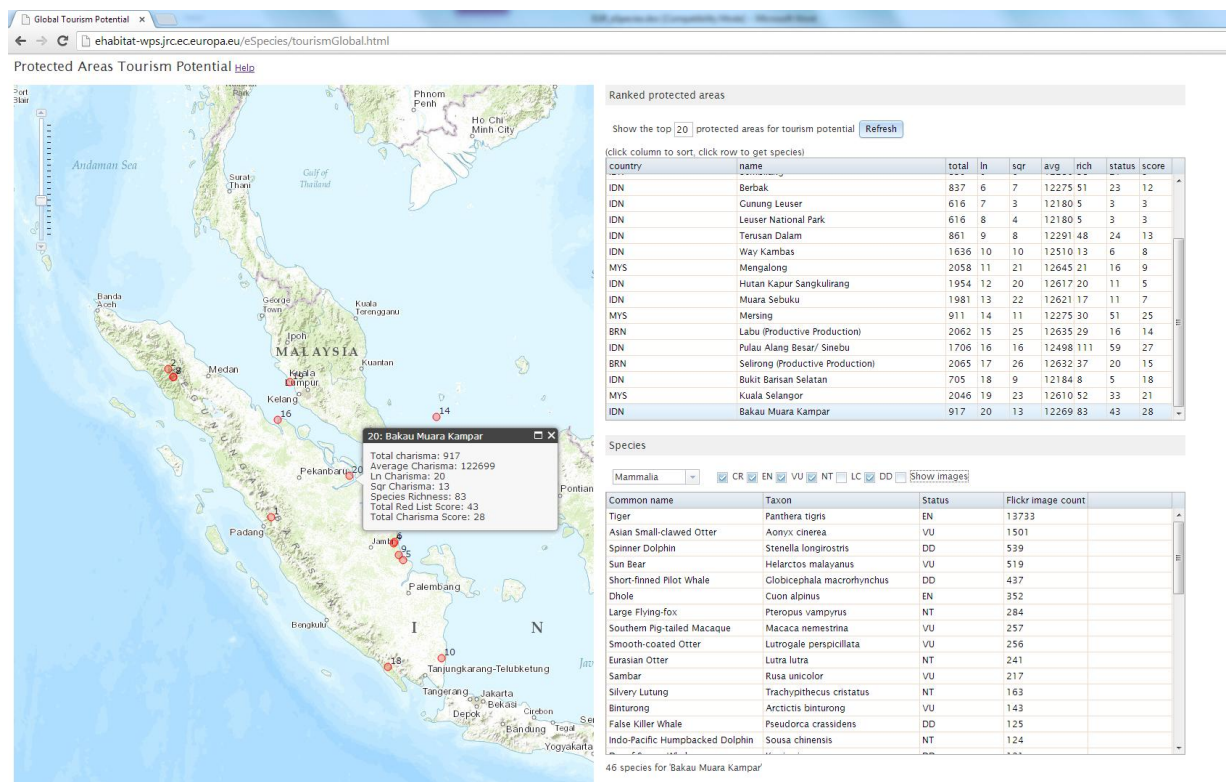


FIGURE 21. THE GLOBAL TOURISM POTENTIAL TOOL IS A TOOL TO HELP IDENTIFY THOSE PROTECTED AREAS WITH THE BEST POTENTIAL FOR TOURISM BASED ON SPECIES CHARISMA.

3.4.2. Integration with Desktop Tools

The REST Web Services can also be consumed by standard consumer Desktop-based software, for example Microsoft Excel. The intention is to make the REST services as widely accessible as possible to ensure their maximum uptake. Using these desktop tools to access the services means that if the underlying data is updated at all the client software will immediately see the updated data. This is particularly important where data is being regularly updated (i.e. the DOPA Validator will allow experts to validate the presence/absence of species in selected areas). Full details of how to access and use the REST services from Desktop tools is given in the REST Services Directory Help.

3.5. Issues

3.5.1. Requirements Specification

One of the biggest challenges with the eSpecies services is in the specification of the products and services. With such a wide range of potential users, use cases and application of the eSpecies products and services it can be difficult to set priorities and scope out the work that will be delivered. Future development work on eSpecies will include logging all user requirements in a catalogue and providing some mechanism for the community of users to democratically prioritise these requirements (see section 4.2.).

3.5.2. Data Issues

Data quality

The species distribution data included in this analysis reflect the current state of knowledge of the geographical distributions of the taxon assessed. They do not, of course, represent all amphibian, mammal and bird species in existence, but instead they are used as indicators of the diversity within that taxon. In addition, the IUCN Red List Data is largely based on expert opinion and the range maps are broad approximations for many species (especially species of least concern). Mapping scales also vary between taxa and species. Amphibians with an extent of occurrence of a few kilometres are likely to be mapped more precisely than a mammal which has a range of thousands of kilometres. Range maps are also often generated as a combination of models and expert consensus and the results inevitably contain uncertainties which can be very large. One will therefore regularly find discrepancies between the theoretical species list and the actual number of species present in a given protected area. It is the objective of the forthcoming DOPA Validator to improve these theoretical lists by inviting experts and park managers to validate these lists for the most significant species and to further provide information on species abundance, information that is currently absent from the DOPA Explorer (Beta), although GBIF data can provide some information.

Another mechanism to provide some feedback to the IUCN Species Specialist groups who are responsible for creating the species range map data is to provide them maps of the intersection of the range maps with land cover data. Where the species habitat requirements are well known the range maps can be improved by excluding areas of unsuitable habitat.

Species data quality becomes a more significant issue at the site level, the level at which conservation implementation actually takes place. For these analyses much finer spatial data are required. For example, the identification of Important Bird Areas (IBAs), developed and promoted by BirdLife International since the early 1980s, has been facilitated by the compilation of locality data for threatened species in, among others, Red Data Books, which subsequently enables 'site-specific synthesis'. This work has subsequently been expanded to include other taxa under the auspices of IUCN and partners to facilitate the identification of key biodiversity areas – sites of known global conservation importance based on confirmed presence of either threatened or irreplaceable biodiversity (where irreplaceable includes restricted-range, congregatory or biome-restricted

species). Finally, a third process to identify sites known to be important will involve access to biodiversity information through GBIF, which would serve to make data available for species that have not yet been assessed under the IUCN process (and could permit independent assessment of IUCN species ranges for some taxa). This would also provide point data to allow an outline assessment of the habitat requirements of species, establishing the potential for ecological niche modelling.

There are also some data quality issues with the WDPA which are currently being addressed. The protected area information that is provided to UNEP-WCMC from the national authorities varies hugely in its quality and coverage but because the data is provided by the national authority it is legally valid. There are issues of incorrect boundaries, missing attribute information (particularly the IUCN management category which indicates how protected the area is) and other data quality issues. However, UNEP-WCMC are working with the national providers to improve the quality and the WDPA is still the most comprehensive database on global protected areas.

Documentation and transparency

Another data issue that is related to data quality is the data documentation. Many of the global biodiversity datasets that are used for eSpecies analyses have been derived over a number of years and in some cases the original source of the data has been lost. In other cases, the processing that has been done on the data is not documented transparently so it can be hard to judge whether a particular dataset is fit for a specific purpose or not.

Completeness

The final main data issue relates to the completeness of the species data. The global RL is not the only source of information on species ranges - it is one of many red list products that are produced within the conservation community. There are also red lists available at the national level and regional level (e.g. the European Red List) and in some cases broken down by taxonomic group, e.g. The Red Data Book of the Mammals of South Africa. All of these products contain valuable information on the status of species, often in more spatial detail than the global RL, but few of them have been incorporated or consulted in the creation of the global RL.

A specific issue with completeness relates to the lack of plant data on the RL. While the main taxonomic groups have been reviewed and spatial data is available there is not yet an assessment for plant species. We are currently exploring the availability of other sources of plant data to supplement the RL.

3.5.3. Collaborative Development

EX-JRC, UNEP-WCMC and IUCN are working together on the DOPA with the joint aims to deliver key reference biodiversity data, but there are many other organizations who are involved in this landscape and coordination is a real issue. The recent work that GBIF has led for the Global Biodiversity Informatics Outlook (GBIO) should help to improve the coordination, the risks to reinvent the wheel being always present. We hope our open data and open source development approach will encourage the reuse of our work. The GEO BON initiative should also further help to coordinate the activities of the organisations concerned.

3.6. Next steps

3.6.1. Improving data quality

Perhaps the most important aspect of the eSpecies that needs to be addressed is the issue of data quality. While many products and services can be used for conservation gain, there are also many others that would benefit from improvements in the data. Future work will concentrate on the DOPA Validator - a tool for community feedback on the eSpecies and other analyses and working with IUCN on improving the content within the RL. A significant initiative that will help to improve

the quality of the species range data is the Map of Life project which aims to integrate a whole suite of species datasets to produce value-added outputs with an assessment of uncertainty.

3.6.2. Additional Biodiversity Value Analyses

The RL has been the main source for information on species distributions within the first part of the eSpecies development, but future work will include the integration of additional datasets from GBIF and include probabilistic species distribution data (i.e. from modelled sources). These data can be used alongside the RL to help derive measures of uncertainty in the range data.

Another area that will be further developed is to analyse the results of the species gap analysis and to identify those areas where there are concentrations of species that have no protection, i.e. a hotspot analysis.

3.6.3. Complementarity Support

Many users within the conservation community who are involved in the site designation process use conservation planning tools, like Marxan for marine areas. It will be especially important to integrate the eSpecies products and services into these conservation planning tools.

3.6.4. Biodiversity Impact Assessment

One of the ambitious aims of the DOPA is to be able to predict the likely impact on biodiversity from a range of environmental and anthropogenic effects, which may be policy driven (e.g. biofuels expansion). These types of scenario modelling analyses will be especially important as policy support for the European Commission and they will become part of a tool called the DOPA Analyst which should be available in 2015.

4. Help and Feedback

End-users of DOPA will find hereafter various ways to get more information about the DOPA Explorer and the eSpecies part of it.

4.1. DOPA Web site

This document is made available from the homepage of DOPA, <http://dopa.jrc.ec.europa.eu/>

This web site is currently the main entry point to access DOPA web sites and web services. It also provides information on latest news, major developments as well as the access to official documentation.

4.2. Engaging with us through the Wiki

Should you be interested in contributing scientifically, technically or as an end-user of DOPA, we would be happy to welcome you on our Wiki to discuss our work at <http://dopa.wikispaces.com/>

4.3. In last resource

You can contact

Grégoire Dubois

European Commission - Joint Research Centre
Institute for Environment and Sustainability
TP 440
Via E. Fermi 2749
I-21027 Ispra (VA)
Italy

Tel : +39 (0)332 786360
Fax : +39 (0)332-789960
Email: gregoire.dubois@jrc.ec.europa.eu

References

- Ball, I.R., H.P. Possingham, and M. Watts (2009). Marxan and relatives: Software for spatial conservation prioritisation. Chapter 14: Pages 185-195 in *Spatial conservation prioritisation: Quantitative methods and computational tools*. Eds Moilanen, A., K.A. Wilson, and H.P. Possingham. Oxford University Press, Oxford, UK.
- Bastin, L., G. Buchanan, A. Beresford, J-F Pekel, G. Dubois (2013). Open-source mapping and services for Web-based land cover validation. *Ecological Informatics*, 14:9-16.
- Brooks, Thomas M., R. A. Mittermeier, G. A. B. da Fonseca, J. Gerlach, M. Hoffmann, J. F. Lamoreux, C. G. Mittermeier, J. D. Pilgrim, A. S. L. Rodrigues (2006). Global biodiversity conservation priorities. *Science*, 313(5783): 58-61.
- Clerici, M., B. Combal, J.F. Pekel, G. Dubois, J. van't Klooster, J.O. Skøien, E. Bartholomé (2013). The eStation, an Earth Observation processing service in support to ecological monitoring. *Ecological Informatics*, (18):162-170.
- Dubois, G., M. Schulz, J. Skøien, A. Cottam, W. Temperley, M. Clerici, E. Drakou, J. van't Klooster, B. Verbeeck, I. Palumbo, P. Derycke, J-F. Pekel, J. Martínez-López, S. Peedell, P. Mayaux (2013). *An introduction to the Digital Observatory for Protected Areas (DOPA) and the DOPA Explorer (Beta)*. EUR 26207 EN, EC. Luxembourg: Publications Office of the European Union 72 p.
- Dubois, G., J. Skøien, M. Schulz, L. Bastin, S. Peedell (2013). eHabitat, a multi-purpose Web Processing Service for ecological modeling. *Environmental Modelling & Software*, 41: 123-133
- EC/COM/2006/0216 final (2006). Communication from the Commission - Halting the loss of biodiversity by 2010 - and beyond - Sustaining ecosystem services for human well-being {SEC(2006) 607} {SEC(2006) 621}
- URL: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52006DC0216:EN:NOT>
- Egoh B., E.G. Drakou, M.B. Dunbar, J. Maes and L. Willemen (2012). *Indicators for mapping ecosystem services: a review*. Report EUR 25456 EN. Luxembourg: Office for Official Publications of the European Communities. 111 pp.
- Gotelli, N. J. and R. K. Colwell (2011). Estimating species richness. in A. E. Magurran and B. J. McGill, editors. *Frontiers in measuring biodiversity*. Oxford University Press, New York. P. 39-54.
- Hartley, A., A. Nelson, P. Mayaux, and JM. Grégoire (2007). *The Assessment of African Protected Areas*. EUR 22780 EN, Luxembourg: Office for Official Publications of the European Communities. 77 pp.
- Hobern, D., Apostolico, A., Arnaud, E., Bello, J. C., Canhos, D., Dubois, G., Field, D., García, E. A., Hardisty, A., et al. (2013). *Global Biodiversity Informatics Outlook: Delivering Biodiversity Knowledge in the Information Age*. GBIF Secretariat. 44 pp.
- Hockings, M (2003). Systems for assessing the effectiveness of management in protected areas. *Bioscience* 53(9), 823(10)
- Maes, J., B. Egoh, L. Willemen, C. Lique, P. Vihervaara, JP. Schägner, B. Grizzetti, E. G. Drakou, A. La Notte, G. Zulian, F. Bouraoui, ML Paracchini, L. Braat, G. Bidoglio (2012). Mapping ecosystem services for policy support and decision making in the European Union. *Ecosystem Services*, 1(1): 31-39.
- Skøien, J., M. Schulz, G. Dubois, I. Fisher, M. Balman, I. May, É. Ó Tuama (2013). Climate change in biomes of Important Bird Areas – results from a WPS application. *Ecological Informatics*, 14:38-43
- UNEP/CBD/COP/10/27, Decision X/7 Examination of the outcome-oriented goals and targets and

associated indicators and consideration of their possible adjustment for the period beyond 2010, Paragraph 2.

UNEP/CBD/COP/11/35, Decision XI/24 Protected Areas, Paragraph 8.

European Commission

EUR 26505 – Joint Research Centre – Institute for Environment and Sustainability

Title: eSpecies. An e-service for mapping species

Authors: Andrew Cottam, Michael Schulz, William Temperley, Grégoire Dubois

Luxembourg: Publications Office of the European Union

2013 – 46 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online), ISSN 1018-5593 (print)

ISBN 978-92-79-35459-5 (pdf)

ISBN 978-92-79-35460-1 (print)

doi: 10.2788/72388

Abstract

The Digital Observatory for Protected Areas (DOPA) is conceived as a set of distributed Critical Biodiversity Informatics Infrastructures (databases, web modelling services, broadcasting services, ...) combined with interoperable web services to provide a large variety of end-users including park managers, decision-makers and researchers with means to assess, monitor and possibly forecast the state and pressure of protected areas at local, regional and global scale.

In particular, the DOPA aims to

- provide the best available material (data, indicators, models) agreed on by contributing institutions which can serve for establishing baselines for research and reporting (i.e. Protected Planet Report, National Biodiversity Strategies and Action Plans, ...);
- provide free analytical tools to support the discovery, access, exchange and execution of web services (databases and modelling) designed to generate the best available material but also for research purposes, decision making and capacity building activities for conservation;
- provide an interoperable and, as much as possible, open source framework to allow institutions to get their own means to assess, monitor and forecast the state and pressure of protected areas and help these to further engage with the organizations hosting critical biodiversity informatics infrastructures.

It is the purpose of this document to introduce the readers to eSpecies, the component of the DOPA providing the services focusing on the delivery of species information and products on a 1 km grid at the global scale. In particular, the readers will find here the necessary instruction to access and use our services as well as some information about the possible uses and limitations of the proposed products and services

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

